FORM PTO-1390 (Modified) (REV 11-98) U.S. DEPARTMENT OF COMMERCE PATENT AND TRADEMARK OFFICE

TRANSMITTAL LETTER TO THE UNITED STATES DESIGNATED/ELECTED OFFICE (DO/EO/US) CONCERNING A FILING UNDER 35 U.S.C. 371

ATTORNEY'S DOCKET NUMBER	U.S. APPLICATION NO. (IF KNOWN, SEE 37 CFR)
MBHB00-615	**09/601852**

INTERNATIONAL APPLICATION NO.	INTERNATIONAL FILING DATE	PRIORITY DATE CLAIMED
PCT/EP99/00860	**09 February 1999**	**10 February 1998**

TITLE OF INVENTION

NOVEL ENDO-XYLOGALACTURONASE

APPLICANT(S) FOR DO/EO/US

1) Petrus J. A. Meeuwsen 2) Cecile J. B. Van Der Vlugt-Bergmans 3) Jean Paul Vincken 4) Gerrit Beldman 5) Alphons G. J. Voragen 6) Margareta A. Herweijer 7) Albert J. J. Van Ooijen

Applicant herewith submits to the United States Designated/Elected Office (DO/EO/US) the following items and other information:

1. ☒ This is a **FIRST** submission of items concerning a filing under 35 U.S.C. 371.
2. ☐ This is a **SECOND** or **SUBSEQUENT** submission of items concerning a filing under 35 U.S.C. 371.
3. ☐ This is an express request to begin national examination procedures (35 U.S.C. 371(f)) at any time rather than delay examination until the expiration of the applicable time limit set in 35 U.S.C. 371(b) and PCT Articles 22 and 39(1).
4. ☒ A proper Demand for International Preliminary Examination was made by the 19th month from the earliest claimed priority date.
5. ☒ A copy of the International Application as filed (35 U.S.C. 371 (c) (2))
   a. ☒ is transmitted herewith (required only if not transmitted by the International Bureau).
   b. ☐ has been transmitted by the International Bureau.
   c. ☐ is not required, as the application was filed in the United States Receiving Office (RO/US).
6. ☐ A translation of the International Application into English (35 U.S.C. 371(c)(2)).
7. ☒ A copy of the International Search Report (PCT/ISA/210).
8. ☐ Amendments to the claims of the International Application under PCT Article 19 (35 U.S.C. 371 (c)(3))
   a. ☐ are transmitted herewith (required only if not transmitted by the International Bureau).
   b. ☐ have been transmitted by the International Bureau.
   c. ☐ have not been made; however, the time limit for making such amendments has NOT expired.
   d. ☐ have not been made and will not be made.
9. ☐ A translation of the amendments to the claims under PCT Article 19 (35 U.S.C. 371(c)(3)).
10. ☐ An oath or declaration of the inventor(s) (35 U.S C. 371 (c)(4)).
11. ☒ A copy of the International Preliminary Examination Report (PCT/IPEA/409).
12. ☐ A translation of the annexes to the International Preliminary Examination Report under PCT Article 36 (35 U.S.C. 371 (c)(5)).

Items 13 to 20 below concern document(s) or information included:

13. ☒ An Information Disclosure Statement under 37 CFR 1.97 and 1.98.
14. ☐ An assignment document for recording. A separate cover sheet in compliance with 37 CFR 3.28 and 3.31 is included.
15. ☒ A **FIRST** preliminary amendment.
16. ☐ A **SECOND** or **SUBSEQUENT** preliminary amendment.
17. ☐ A substitute specification.
18. ☐ A change of power of attorney and/or address letter.
19. ☐ Certificate of Mailing by Express Mail
20. ☒ Other items or information:

Sequence Listing in computer readable form
Acknowledgement Postcard

U.S. APPLICATION NO. (IF KNOWN, SEE 37 CFR)	INTERNATIONAL APPLICATION NO.	ATTORNEY'S DOCKET NUMBER
09/601852	PCT/EP99/00860	MBHB00-615

21. The following fees are submitted:.

BASIC NATIONAL FEE (37 CFR 1.492 (a) (1) - (5)) :

		CALCULATIONS PTO USE ONLY
☒	Neither international preliminary examination fee (37 CFR 1.482) nor international search fee (37 CFR 1.445(a)(2) paid to USPTO and International Search Report not prepared by the EPO or JPO	**$970.00**
☐	International preliminary examination fee (37 CFR 1.482) not paid to USPTO but Internation Search Report prepared by the EPO or JPO	**$840.00**
☐	International preliminary examination fee (37 CFR 1.482) not paid to USPTO but international search fee (37 CFR 1.445(a)(2)) paid to USPTO	**$690.00**
☐	International preliminary examination fee paid to USPTO (37 CFR 1.482) but all claims did not satisfy provisions of PCT Article 33(1)-(4)	**$670.00**
☐	International preliminary examination fee paid to USPTO (37 CFR 1.482) and all claims satisfied provisions of PCT Article 33(1)-(4)	**$96.00**

ENTER APPROPRIATE BASIC FEE AMOUNT = **$970.00**

Surcharge of **$130.00** for furnishing the oath or declaration later than ☐ 20 ☒ 30 months from the earliest claimed priority date (37 CFR 1.492 (e)). **$130.00**

CLAIMS	NUMBER FILED	NUMBER EXTRA	RATE	
Total claims	46 - 20 =	26	x $18.00	**$468.00**
Independent claims	5 - 3 =	2	x $78.00	**$156.00**
Multiple Dependent Claims **(check if applicable).**			☒	**$260.00**

TOTAL OF ABOVE CALCULATIONS = **$1,984.00**

Reduction of 1/2 for filing by small entity, if applicable. Verified Small Entity Statement must also be filed (Note 37 CFR 1.9, 1.27, 1.28) **(check if applicable).** ☐ **$0.00**

SUBTOTAL = **$1,984.00**

Processing fee of **$130.00** for furnishing the English translation later than ☐ 20 ☐ 30 months from the earliest claimed priority date (37 CFR 1.492 (f)). + **$0.00**

TOTAL NATIONAL FEE = **$1,984.00**

Fee for recording the enclosed assignment (37 CFR 1.21(h)). The assignment must be accompanied by an appropriate cover sheet (37 CFR 3.28, 3.31) **(check if applicable).** ☐ **$0.00**

TOTAL FEES ENCLOSED = **$1,984.00**

Amount to be: refunded	$
charged	$

☒ A check in the amount of **$1,984.00** to cover the above fees is enclosed.

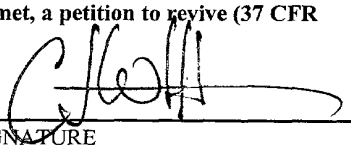
☐ Please charge my Deposit Account No. in the amount of to cover the above fees.
A duplicate copy of this sheet is enclosed.

☒ The Commissioner is hereby authorized to charge any fees which may be required, or credit any overpayment to Deposit Account No. **13-2490** A duplicate copy of this sheet is enclosed.

NOTE: Where an appropriate time limit under 37 CFR 1.494 or 1.495 has not been met, a petition to revive (37 CFR 1.137(a) or (b)) must be filed and granted to restore the application to pending status.

SEND ALL CORRESPONDENCE TO:

Curt J. Whitenack
MCDONNELL BOEHENEN HULBERT & BERGHOFF
300 South Wacker Drive
Suite 3200
Chicago, Illinois 60606
USA



SIGNATURE

Curt J. Whitenack
NAME

36,054
REGISTRATION NUMBER

09 August 2000
DATE

UNITED STATES PATENT AND TRADEMARK OFFICE
(Case No. MBHB00-615)

PATENT

In the Application of:	
Petrus Johanne Albertus Meeuwsen, et al.	**Group Art Unit: Unassigned**
U.S. National Phase of International Application No. PCT/EP99/00860	**Examiner: Unassigned**
Filed: 09 February 1999	
For: Novel Endo-Xylogalacturonase	

PRELIMINARY AMENDMENT

Assistant Commissioner for Patents
Washington, D.C. 20231

Dear Sir:

Prior to the examination of the above-identified patent application, please enter the following amendments and consider the following remarks.

IN THE CLAIMS:

Please amend claims 4, 7, 10, 12 through 20, 22, 25, 26, 28, 30, 31, 33, 34 and claim 38 as follows:

4. (Amended) A polypeptide according to [any preceding] claim 1, which comprises the sequence set out in SEQ ID No. 2, or a sequence substantially homologous thereto, or a fragment of either sequence.

7. (Amended) A polynucleotide encoding a polypeptide according [to any one of the preceding] claim[s] 1.

10. (Amended) An isolated polynucleotide according to claim 7[, 8 or 9] obtainable from a fungus.

12. (Amended) A polynucleotide probe which comprises a fragment of at least 15 nucleotides of a polynucleotide as defined in [any of] claim[s] 7 [to 11].

13. (Amended) a vector comprising a polynucleotide as defined in [any one of] claim[s] 7 [to 12].

14. (Amended) An expression vector comprising a polynucleotide as defined in [any one of] claim[s] 7 [to 11] operably linked to one or more regulatory sequences capable of directing expression of the polynucleotide in a host cell.

15. (Amended) A host cell transformed or transfected with, comprising or incorporating a vector according to [any one] of claim[s] 13 [to 14].

16. (Amended) A host cell comprising or harboring a polynucleotide according to [any one of] claim[s] 7 [to 11] wherein the polynucleotide is heterologous to the genome of the host cell.

17. (Amended) A host cell according to claim 15 [or claim 16] which is a yeast cell.

18. (Amended) A method for producing a polypeptide according to [any one of] claim[s] 1 [to 6] which comprises incubating or culturing a host cell according to [any one of] claim[s] 15 [to 17] under conditions which allow the expression of the polypeptide, and optionally purifying the polypeptide.

19. (Amended) A host cell comprising or expressing a polypeptide according to [any one of] claim[s] 1 [to 6] therein the polypeptide is heterologous to the host cell.

20. (Amended) A composition comprising a polypeptide according to [any one of] claim[s] 1 [to 6].

22. (Amended) A method of treating a plant material, the method comprising contacting the plant material with a polypeptide according to [any one of] claim[s] 1 [to 6 or a composition according to claim 20 or claim 21].

25. (Amended) A method according to claim 22 [or 23] wherein the treatment comprises endo-type cleaving of xylogalacturonan subunits of a pectin component of the material.

26. (Amended) A method according to [any of] claim[s] 22 [to 24] wherein the material comprises a plant, plant pulp, plant extract or an edible foodstuff or ingredient therefore.

28. (Amended) A processed plant material obtainable by contacting a plant material with a polypeptide according to [any one of] claim[s] 1 [to 6 or a composition according to claims 20 or claim 21, or which results from a method according to any of claim 22 to 26].

30. (Amended) A method for reducing the viscosity of a plant material, the method comprising contacting the plant material with a polypeptide according to [any one of] claim[s] 1 [to 6 or a composition according to claim 20 or claim 21 in an amount and under conditions effective to degrade pectin contained in the material].

31. (Amended) Use of a polypeptide according to [any one of] claim[s] 1 [to 6 or a composition according to claim 20 or claim 21] in a method of treating plant material.

33. (Amended) Use of a polypeptide according to [any one of] claim[s] 1 [to 6 or a composition according to claim 20 or claim 21] in a method of processing plant pulp, juice or extract which method comprises incubating the pulp, juice or extract with the polypeptide or composition to at least partially degrade pectin.

34. (Amended) An (animal) feed or foodstuff comprising a polypeptide according to [any one of the] claim[s] 1 [to 6].

38. (Amended) An assay according to claim 35 [or 36] which comprises measuring the amount of Cu(II) reduced to Cu(I) by the carbohydrates, optionally by contact with bicinchoninic acid (BCA) and determining the amount of BCA-Cu (I) complex formed.

Please add new claims 39-44 as follows:

39. A method of treating a plant material, the method comprising contacting the plant material with a composition according to claim 20.

40. A processed plant material obtainable by contacting a plant material with a composition according to claim 20.

41. A processed plant material obtainable by contacting a plant material with a composition which results from a method according to any of claim 22.

42. A method for reducing the viscosity of a plant material, the method comprising contacting the plant material with a composition according to claim 20 in an amount and under conditions effective to degrade pectin contained in the material.

43. Use of a composition according to claim 20 in a method of treating plant material.

44. Use of a composition according to claim 20 in a method of processing plant pulp, juice or extract which method comprises incubating the pulp, juice or extract with the polypeptide or composition to at least partially degrade pectin.

REMARKS

Applicants respectfully request the Examiner to enter this Preliminary Amendment. Claims 4, 7, 10, 12 through 20, 22, 25, 26, 28, 30, 31, 33, 34 and 38 have been amended to remove muliple dependancies and more particulary point out the subject matter Applicants consider the invention. The new claims meet the requirements of 35 U.S.C. § 112. No new matter has been added by these amendments. Support for the amendments and new claims is found in the specification as originally filed.

With these amendments, the claims pending are 1-44.

Applicants respectfully solicit allowance of the claims as amended and passage of the case to issuance. If in the opinion of the Examiner a telephone conference would expedite the prosecution of the subject application, the Examiner is invited to call the undersigned attorney at (312) 913-0001

Respectfully submitted,

Curt J. Whitenack, Registration No. 36,054
One of the Attorneys for Applicant(s)
McDonnell, Boehnen, Hulbert & Berghoff
300 South Wacker Drive
Chicago, Illinois 60606
(312) 913-0001 (telephone)
(312) 913-0003 (facsimile)

Date: August 9, 2000

NOVEL ENDO-XYLOGALACTURONASE

Field of the Invention

The present invention relates to a novel endo-xylogalacturonase (XGH) and homologues thereof. It further relates to the use of the endo-xylogalacturonase in a method of processing plant or pectin-containing material to produce fruit juice and other plant extracts.

Background to the Invention

Enzyme preparations are often used during the processing of plant materials, for example in the steps of extraction and liquefaction of fruit and fruit juice and their filtration and clarification. Commercial enzyme preparations contain a mixture of enzymes which degrade the pectin polymers which are a major component of plant cell walls. Such enzymes include pectin lyases, polygalacturonases, pectin esterases, celluloses, xyloglucanases, galactanases and arabinanases.

Pectins occur in nature as constituents of higher plant cell walls. They are found in primary cell wall lamella where they are embedded in between the cellulose fibrils. The composition of pectin is variable among plant species and moreover dependent on the age and the maturity of the fruit. Among the richer sources of pectins are lemons and oranges, which can represent up to 30% of polysaccharides present.

Most pectin polymers are comprised of 'smooth' homogalacturonan regions and ramified 'hairy' regions. The 'smooth' regions consist of a linear homogalacturonan backbone. The 'hairy' regions of apples consists of three different subunits: subunit I is xylogalacturonan (a galacturonan backbone heavily substituted with xylose); subunit II is a short section of a rhamnogalacturonan backbone, rich in relatively long arabinan, galactan and/or arabinogalactan side chains (the 'hairs'); and subunit III is a rhamnogalacturonan oligomer, having a backbone consisting of an alternating sequence of rhamnose and galacturonic acid residues.

Many of the well-known pectinases used in industrial food processing degrade only the 'smooth' part of the pectin polymer leaving the 'hairy' regions intact. Consequently, during for example apple juice production, the non-degraded parts of the pectin polymer cause production losses due to inefficient filtration as a result of fouling of the (ultra)filtration membrane.

Several enzymes have been reported which can degrade parts of the 'hairy' region, for example the rhamnogalacturonan regions of the backbone (subunit III). These enzymes are referred to as rhamnogalacturonases (RGases), of which there are several types. However, so far the xylogalacturonan part of the 'hairy' regions (subunit I) has been resistant to enzymatic digestion and so in prior art enzymatic endo-digestion processes the xylogalacturonan is left as an inert carbohydrate.

Since xylogalacturonan has also been found in many other plants, e.g. leguminous plants like soybeans and peas, watermelons, grapes and pine pollen, enzymes to degrade this polymer would be useful for the processing of plant material.

An exo-galacturonase (42kDa, SDS-PAGE) has been identified[1] that is not hindered by the single unit xylose side-chains and is able to degrade xylogalacturonan using a soluble 'hairy' pectic polysaccharide from soy as substrate. This enzyme acts in an exo-fashion as it yields galacturonic acid or a disaccharide consisting of galacturonic acid and xylose. The enzyme was purified to near homogeneity (Fractions HTP2 and Q2) and partially characterized. By contrast to known RGases (which do not degrade homogalacturonic acid) this enzyme is not very specific for xylogalacturonan as it also acts on pectic acid. In addition, this enzyme is not able to digest the xylogalacturonan backbone in a random fashion, and therefore to date there are no known enzymes possessing endo-xylogalacturonase activity.

Disclosure of the Invention

The present invention has resulted from the isolation and characterization of a novel endo-xylogalacturonase and cDNA encoding it. The endo-xylogalacturonase cDNA sequence is set out in SEQ. ID No. 1. The amino acid sequence of the ORF from nucleotides 98 to 1315 is set out in SEQ. ID No. 2.

In a first aspect of the invention there is provided an (e.g. isolated and/or purified)

polypeptide possessing endo-xylogalacturonase activity. There is also provided a polypeptide comprising an endo-xylogalacturonase, such as a polypeptide comprising the sequence set out in SEQ ID No. 2, or a polypeptide substantially homologous thereto, or a fragment of the polypeptide of SEQ ID No. 2 having at least 5 amino acids.

The polypeptide of the invention preferably has one or more of the following additional features, namely it.

(1) possesses endo-xylogalacturonase activity;

(2) has an optimal pH-range of from 2.5 to 6;

(3) has optimum activity at a temperature of from 50 to 70°C; and/or

(4) has a molecular weight (deglycosylated) of from 40 to 50 kDa.

"Endo-xylogalacturonase activity" is defined as the ability to cleave a galacturonic acid polymer (for example as found in pectin) which may be at least partially substituted with xylose at internal glycosidic bonds. The activity thus allows cleavage between adjacent galacturonan non-terminal units (where neither of such units is at the end of the polymer, which is in contrast to exo activity where the end unit would be cleaved). Preferably the cleavage occurs at a [galacturonic acid (1-4) galacturonic acid] linkage. Preferably, the polypeptide does not cleave terminal xylose residues from xylose substituted galacturonic acid residues, for example a [galacturonic acid (3-1) xylose] linkage. The polypeptide may preferentially cleave in between two adjacent non-xylose substituted galacturonan units. The substrate polymer may be from 40 to 80% (e.g. xylose) substituted.

The two galacturonic acid residues between which the polypeptides of the invention cleave may both be (xylose) substituted, or only one may be (xylose) substituted or (preferably) neither may be (xylose) substituted. Alternatively or in addition the two galacturonic acid residues may both be methylated, or one may be methylated, or (preferably) neither may be methylated.

Preferably, the polypeptide of the invention is obtainable from a microorganism which possesses a gene encoding an enzyme with endo-xylogalacturonase activity. More preferably the microorganism is a microbial organism, preferably fungal, and optimally a filamentous fungi. Preferred organisms are thus of the genera *Aspergillus*, *Trichoderma*, *Penicillium*, *Acremonium*, *Fusarium*, *Humicola*, *Neurospora*, *Mucor*, *Scytallidium*, *Myceliophtora*, *Thielavia*, *Talaromyces*, *Thermomyces*, *Thermoascus*, *Chaetomium*,

Sporotrichum, *Corynascus*, *Calcarisporiella* or *Mycelia*. Optionally the organism is of the species from the *Aspergillus niger* group (as defined by Raper and Fennell, The Genus *Aspergillus*, The Williams & Wilkins Company, Baltimore, pp 293-344, 1965), specifically including but not limited to *Aspergillus niger*, *Aspergillus awamori*, *Aspergillus tubigensis*, *Aspergillus aculeatus*, *Aspergillus foetidus*, *Aspergillus japonicus* or *Aspergillus ficuum*.

In a second aspect, the present invention provides an (e.g. isolated and/or purified) polynucleotide encoding a polypeptide of the first aspect of the invention. For example the present invention provides a polynucleotide encoding an endo-xylogalacturonase, such an endo-xylogalacturonase whose amino acid sequence is set out in SEQ ID No. 2. The present invention further provides a polynucleotide encoding a polypeptide having substantial amino acid sequence homology to the amino acid sequence set out in SEQ ID No. 2. Also provided is a polynucleotide selected from:

(a) polynucleotides comprising the nucleotide sequence set out in SEQ ID No. 1, or the complement thereof;

(b) polynucleotides comprising a nucleotide sequence capable of hybridising to the nucleotide sequence set out in SEQ ID No. 1, or a fragment thereof;

(c) polynucleotides comprising a nucleotide sequence capable of hybridising to the complement of the nucleotide sequence set out in SEQ ID No. 1, or a fragment thereof; and/or

(d) polynucleotides comprising a polynucleotide sequence which is degenerate as a result of the genetic code to the polynucleotides defined in (a), (b) or (c).

A polynucleotide of the invention also includes a polynucleotide which:

a. encodes a polypeptide having endo-xylogalacturonase activity, which polynucleotide is:

(1) the coding sequence of SEQ ID No. 1;

(2) a sequence which hybridises selectively to the complement of sequence defined in (1); or

(3) a sequence that is degenerate as a result of the genetic code with respect to a sequence defined in (1) or (2); or

b. is a sequence complementary to a polynucleotide defined in (a).

The term "capable of hybridizing" means that the target polynucleotide of the

invention can hybridize to the nucleic acid used as a probe (for example the nucleotide sequence set out in SEQ. ID No.1, or a fragment thereof or the complement thereof) at a level significantly above background. The background hybridization may occur because of, for example, other polynucleotides, such as DNA, present in, for example a cDNA/genomic library being screened. In this event, background implies a level of signal generated by interaction between the probe and a non-specific polynucleotide member of the library which is less than 10 fold, preferably less than 100 fold, as intense as the specific interaction observed with the target polynucleotide. The intensity of interaction may be measured, for example, by radiolabelling the probe, e.g. with ^{32}P. Suitable conditions are described later.

Preferably, the polynucleotide of the invention is obtainable from the same organism as the polypeptide, such as a fungus, in particular a fungus of the genus *Aspergillus*.

The present invention also provides a polynucleotide probe which comprises a fragment of at least 15 nucleotides of a polynucleotide of the invention as described above.

In a third aspect, the invention provides vectors comprising a polynucleotide of the invention, including cloning and expression vectors, and in a fourth aspect methods of growing, transforming or transfecting such vectors in a suitable host cell, for example under conditions in which expression of a polypeptide of, or encoded by a sequence of, the invention occurs. Provided in a fifth aspect are host cells comprising a polynucleotide or vector of the invention wherein the polynucleotide is heterologous to the genome of the host cell. The term "heterologous to the genome of the host cell" means that the polynucleotide does not naturally occur in the genome of the host cell. Preferably, the host cell is a yeast cell, for example a yeast cell of the genus *Kluyveromyces* or *Saccharomyces* or a fungal cell, for example of the genus *Aspergillus*.

The polypeptides of the invention which possess endo-xylogalacturonase activity may be used in a sixth aspect to treat plant material including plant pulp and plant extracts. For example, they may be used to treat apple pulp and/or raw juice during the production of apple juice. Conveniently the polypeptide of the invention is combined with suitable carriers or diluents including buffers to produce a composition/ enzyme preparation. Thus the present invention provides in a seventh aspect a composition comprising a polypeptide of the invention. The composition may further comprise additional ingredients such as one or more enzymes, for example pectinases, including endo-arabinanase and

rhamnogalacturonase, cellulases and/or xyloglucanases.

The polypeptides and compositions of the invention may therefore be used in a method of processing plant material to degrade or modify the pectin constituents of the cell walls of the plant material. Thus in an eighth aspect, the present invention provides a method of degrading or modifying a plant cell wall which method comprises contacting the plant cell wall with a polypeptide or composition of the invention.

The invention also provides a method of processing a plant material which method comprises contacting the plant material with a polypeptide or composition of the invention to degrade or modify the pectin in the plant material Preferably the plant material is a plant pulp or plant extract.

In particular, the degradation preferably comprises endo-type cleaving of xylogalacturonan subunits of a pectin component of the plant cell wall. The plant material is preferably a fruit or vegetable pulp or fruit or vegetable extract, for example apple pulp or apple juice.

The present invention further provides a processed plant material obtainable by contacting a plant material with a polypeptide or composition of the invention. Preferably the processed plant material is a fruit or vegetable juice, for example apple juice.

The present invention also provides a method for reducing the viscosity of a plant extract which method comprises contacting the plant extract with a polypeptide or composition of the invention in an amount effective in degrading pectins contained in said plant extract.

Preferred features and characteristics of one aspect of the invention are applicable to another aspect *mutatis mutandis*.

Detailed description of the invention

A. Polynucleotides.

The invention provides a polynucleotide which.

a. encodes a polypeptide that has endo-xylogalacturonase activity, which polynucleotide is:

(1) the coding sequence of SEQ ID No. 1;

(2) a sequence that hybridizes selectively to the complement of the sequence

defined in (1); or

(3) a sequence that is degenerate as a result of the genetic code with respect to the nucleic acid sequence defined in (1) or (2); or

b. is a sequence complementary to a polynucleotide defined in (a).

The polynucleotides of the invention also include variants of the coding sequence of SEQ ID No. 1 which have endo-xylogalacturonase activity. Variants may be formed by additions, substitutions and/or deletions. Such variants may thus have the ability to cleave internally a galacturonic acid polymer. Typically a polynucleotide of the invention comprises a continuous sequence of nucleotides which is capable of hybridizing under selective conditions to the complement of the coding sequence of SEQ ID No. 1.

A polynucleotide of the invention and complement of the coding sequence of SEQ ID No. 1 can hybridize at a level significantly above background. Background hybridization may occur, for example, because of other cDNA's present in a cDNA library. The signal level generated by the interaction between a polynucleotide of the invention and the complement of the coding sequence of SEQ ID No. 1 is typically at least 10-fold, preferably at least 100-fold, as intense as interactions between other polynucleotides and the coding sequence of SEQ ID No. 1. The intensity of interaction may be measured, for example, by radiolabelling the probe, for example with ^{32}P. Selective hybridization may typically be achieved using conditions of low stringency (for example, 0.03M sodium chloride and 0.03M sodium citrate at about 40°C), medium stringency (for example, 0.03M sodium chloride and 0.03M sodium citrate at about 50°C) or high stringency (for example, 0.03M sodium chloride and 0.03M sodium citrate at about 60°C).

A preferred polynucleotide is capable of selectively hybridizing to complement the DNA sequence of SEQ ID No. 1 will generally have at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 98% or at least 99% sequence identity to the coding sequence of SEQ ID No. 1 over a region of at least 20, preferably at least 30, for instance at least 40, at least 60, or preferably at least 100 continuous nucleotides or most preferably over the full length of SEQ ID No. 1.

Any combination of the above mentioned degrees of sequence identity and minimum sizes may be used to define polynucleotides of the invention, with the more stringent combinations (that is to say higher sequence identity over longer lengths) being preferred. Thus, for example, a polynucleotide which has least 90% sequence identity over 25,

preferably over 30 nucleotides, is preferred, as is a polynucleotide which has at least 95% sequence identity over 40 nucleotides.

The coding sequence of SEQ ID No. 1 may be modified by nucleotide substitutions, for example from 1, 2 or 3 to 10, 25, 50 or 100 substitutions. The polynucleotide of SEQ ID No. 1 may alternatively or additionally be modified by one or more insertions and/or deletions (such as the same number mentioned for substitutions) and/or by an extension to either or both ends. The modified polynucleotide in general encodes a polypeptide which has endo-xylogalacturonase activity. Degenerate substitution may be made and/or substitutions may be made which would result in a conservative amino acid substitution when the modified sequence is translated, for example as shown in the Table on page 12 in the section concerning polypeptides.

Polynucleotides of the invention may comprise DNA or RNA. They may be single or double stranded. They may also be polynucleotides which include within them synthetic or modified nucleotides. A number of different types of modifications to polynucleotides are known in the art. These include a methylphosphonate and phosphorothioate backbones, and addition of acridine or polylysine chains at the 3' and/or 5' ends of the molecule. For the purposes of the present invention, it is to be understood that the polynucleotides described herein may be modified by any method available in the art.

It is to be understood that skilled persons may, using routine techniques, make nucleotide substitutions that do not affect the polypeptide sequence encoded by the polynucleotides of the invention to reflect the codon usage of any particular host organism in which the polypeptides of the invention are to be expressed.

Polynucleotides of the invention may be used as a primer, e.g. a PCR primer, a primer for an alternative amplification reaction, a probe e.g. labelled with a revealing label by conventional means using radioactive or non-radioactive labels, or the polynucleotides may be cloned into vectors.

Such primers, probes and other fragments will be at least 15, preferably at least 20, for example at least 25, 30 or 40 nucleotides in length. There will typically be up to 40, 50, 60, 70, 100 or 150 nucleotides in length. Probes and fragments can be longer than 150 nucleotides in length, for example up to 200, 300, 400, 500, 600, 700 nucleotides in length, or even up to a few nucleotides (such as 5 or 10 nucleotides) short of the coding sequence of SEQ ID No. 1.

Polynucleotides such as a DNA polynucleotide and primers according to the invention may be produced recombinantly, synthetically, or by any means available to those of skill in the art. They may also be cloned by standard techniques. The polynucleotides are typically provided in isolated and/or purified form.

In general, primers will be produced by synthetic means, involving a step-wise manufacture of the desired nucleic acid sequence one nucleotide at a time. Techniques for accomplishing this using automated techniques are readily available in the art.

Longer polynucleotides will generally be produced using recombinant means, for example using a PCR (polymerase chain reaction) cloning techniques. This will involve making a pair of primers (e.g. of about 15-30 nucleotides) to a region of the endo-xylogalacturonase gene which it is desired to clone, bringing the primers into contact with mRNA or cDNA obtained from a fungal, yeast, bacterial plant or prokaryotic cell, performing a polymerase chain reaction under conditions which bring about amplification of the desired region, isolating the amplified fragment (e.g. by purifying the reaction mixture on an agarose gel) and recovering the amplified DNA. The primers may be designed to contain suitable restriction enzyme recognition sites so that the amplified DNA can be cloned into a suitable cloning vector.

Such techniques may be used to obtain all or part of the endo-xylogalacturonase sequence described herein. Genomic clones corresponding to the cDNA of SEQ ID No. 1 or the endo-xylogalacturonase gene containing, for example, introns and promoter regions are within the invention also and may also be obtained in an analogous manner (e.g. recombinant means, PCR, cloning techniques), starting with genomic DNA from a fungal, yeast, bacterial plant or prokaryotic cell.

Although in general the techniques mentioned herein are well known in the art, reference may be made in particular to Sambrook *et al.*, Molecular Cloning, A Laboratory Manual (1989) and Ausubel *et al.*, Current Protocols in Molecular Biology (1995), John Wiley & Sons, Inc.

Polynucleotides which do not have 100% identity with SEQ ID No. 1 but fall within the scope of the invention can be obtained in a number of ways. Thus variants of the endo-xylogalacturonase sequence described herein may be obtained for example by probing genomic DNA libraries made from a range of organisms, for example those discussed as sources of the polypeptides of the invention. In addition, other fungal, plant

or prokaryotic homologues of endo-xylogalacturonase may be obtained and such homologues and fragments thereof in general will be capable of hybridising to SEQ ID No. 1. Such sequences may be obtained by probing cDNA libraries or genomic DNA libraries from other species, and probing such libraries with probes comprising all or part of SEQ ID. 1 under conditions of medium to high stringency (for example 0.03M sodium chloride and 0.03M sodium citrate at from 50°C to 60°C). Nucleic acid probes comprising all or part of SEQ ID No. 1 may be used to probe cDNA libraries from other species, such as those described as sources for the polypeptides of the invention .

Species homologues may also be obtained using degenerate PCR which will use primers designed to target sequences within the variants and homologues encoding conserved amino acid sequences. The primers will contain one or more degenerate positions and will be used at stringency conditions lower than those used for cloning sequences with single sequence primers against known sequences.

Alternatively, such polynucleotides may be obtained by site directed mutagenesis of the endo-xylogalacturonase sequences or variants thereof. This may be useful where for example silent codon changes are required to sequences to optimise codon preferences for a particular host cell in which the polynucleotide sequences are being expressed. Other sequence changes may be desired in order to introduce restriction enzyme recognition sites, or to alter the property or function of the polypeptides encoded by the polynucleotides.

The invention includes double stranded polynucleotides comprising a polynucleotide of the invention and its complement.

Polynucleotides or primers of the invention may carry a revealing label. Suitable labels include radioisotopes such as ^{32}P or ^{35}S, enzyme labels, or other protein labels such as biotin and DIG-hapten. Such labels may be added to polynucleotides or primers of the invention and may be detected using by techniques known *per se*.

The present invention also provides polynucleotides encoding the polypeptides of the invention described below. Since such polynucleotides will be useful as sequences for recombinant production of polypeptides of the invention, it is not necessary for them to be capable of hybridising to the sequence of SEQ ID No. 1, although this will generally be desirable. Otherwise, such polynucleotides may be labelled, used, and made as described above if desired.

B Polypeptides.

A polypeptide of the invention comprises the amino acid sequence set out in SEQ ID No. 2 or a substantially homologous sequence, or a fragment of either sequence and can have endo-xylogalacturonase activity. In general, the naturally occurring amino acid sequence shown in SEQ ID No. 2 is preferred.

In particular, the polypeptide of the invention may comprise:

a. the polypeptide sequence of SEQ ID No. 2;

b. a naturally occurring variant or species homologue thereof; or

c. a protein with at least 60, at least 70, at least 80, at least 90, at least 95, at least 98 or at least 99% sequence identity to (a) or (b).

A variant will be one that occurs naturally, for example in fungal, bacteria, yeast or plant cells and which can function in a substantially similar manner to the protein of SEQ ID No. 2, for example it has endo-xylogalacturonase activity. Similarly a species homologue of the protein will be the equivalent protein which occurs naturally in another species and which can function as an endo-xylogalacturonase enzyme.

Variants and species homology can be obtained by following the procedures described herein for the production of the polypeptide of SEQ ID No. 2 and performing such procedures on a suitable cell source, for example a bacterial, yeast, fungal or plant cell. It will also be possible to use a probe as defined above to probe libraries made from yeast, bacterial, fungal or plant cells in order to obtain clones including the variants or species homology. The clones can be manipulated by conventional techniques to generate a polypeptide of the invention which can then be produced by recombinant or synthetic techniques known *per se*.

The polypeptide of the invention preferably has at least 60% sequence identity to the protein of SEQ ID No. 2, more preferably at least 70%, at least 80%, at least 90%, at least 95%, at least 97% or at least 99% sequence identity thereto over a region of at least 20, preferably at least 30, for instance at least 40, at least 60, at least 100, 200 or 300 contiguous amino acids or over the full length of SEQ ID No. 2.

The sequence of the polypeptide of SEQ ID No. 2 and of variants and species homologues can thus be modified to provide polypeptides of the invention. Amino acid substitutions may be made, for example from 1, 2 or 3 to 10, 20 to 30 substitutions. The

same number of deletions and insertions may also be made. The modified polypeptide generally retains activity as an endo-xylogalacturonase.

Conserved substitutions may be made according to the following Table, where amino acids in the same block in the second column and preferably in the same line in the third column may be substituted for each other:

ALIPHATIC	Non-polar	G A P
		I L V
	Polar - uncharged	C S T M
		N Q
	Polar - charged	D E
		K R
AROMATIC		H F W Y
OTHER		N Q D E

Polypeptides of the invention also include fragments of the above mentioned full length polypeptides and of variants thereof, including fragments of the sequence set out in SEQ ID No. 2. Such fragments typically retain activity as an endo-xylogalacturonase. Fragments may be at least 10, 15, 20, 30, 50, 100 or 200 amino acids long.

Polypeptides of the invention may be in a substantially isolated form. It will be understood that the polypeptide may be mixed with carriers or diluents which will not interfere with the intended purpose of the polypeptide and still be regarded as substantially isolated. A polypeptide of the invention may also be in a substantially purified form, in which case it will generally comprise the polypeptide in a preparation in which more than 50%, e.g. more than 80%, 90%, 95%, 98% or 99% by weight of the polypeptide in the preparation is a polypeptide of the invention. Polypeptides of the invention may be chemically modified, for example post-transnationally modified. For example, they may be glycosylated (one or more times) or comprise one or more modified amino acid residues

They may be modified for example by the addition of histidine residues or a T7 tag to assist their identification or purification or by the addition of a signal sequence to promote their secretion from a cell, as discussed below.

Polypeptides of the invention can if necessary be produced by synthetic means although usually they will be made recombinantly as described below.

Particularly preferred polypeptides of the invention include a polypeptide consisting of amino acids 19 to 406 of the amino acid sequence set out in SED. ID No. 2 since this lacks the N-terminal signal peptide which consists of amino acids 1 to 18 of the amino acid sequence of SEQ ID No. 2. The polypeptides and fragments thereof may contain amino acid alterations as defined above.

The use of yeast and fungal host cells is expected to provide for such post-translational modifications (e.g. proteolytic processing, myristilation, glycosylation, truncation, and tyrosine, serine or threonine phosphorylation) as may be needed to confer optimal biological activity on recombinant expression products of the invention.

C. Recombinant Aspects.

Polynucleotides of the invention can be incorporated into a recombinant replicable vector, for example a cloning or expression vector. The vector may be used to replicate the nucleic acid in a compatible host cell. Thus in a further embodiment, the invention provides a method of making polynucleotides of the invention by introducing a polynucleotide of the invention into a replicable vector, introducing the vector into a compatible host cell, and growing the host cell under conditions which bring about replication of the vector. The vector may be recovered from the host cell. Suitable host cells are described below in connection with expression vectors.

Expression Vectors

Preferably, a polynucleotide of the invention in a vector is operably linked to a regulatory sequence which is capable of providing for the expression of the coding sequence by the host cell, i.e. the vector is an expression vector. The term "operably linked" refers to a juxtaposition wherein the components described are in a relationship permitting them to function in their intended manner. A regulatory sequence such as a promoter, enhancer or other expression regulation signal "operably linked" to a coding sequence is positioned in such a way that expression of the coding sequence is achieved under condition compatible with the control sequences.

The vectors may be for example, plasmid, virus or phage vectors provided with an origin of replication, optionally a promoter for the expression of the polynucleotide and optionally a regulation of the promoter.

The DNA sequence encoding the polypeptide is preferably introduced into a suitable host as part of an expression construct in which the DNA sequence is operably linked to expression signals which are capable of directing expression of the DNA sequence in the host cells. For transformation of the suitable host with the expression construct transformation procedures are available which are well known to the skilled person[3,4]. The expression construct can be used for transformation of the host as part of a vector carrying a selectable marker, or the expression construct is co-transformed as a separate molecule together with the vector carrying a selectable marker. The vectors may contain one or more selectable marker genes.

Preferred selectable markers[3,4] include but are not limited to e.g. versatile marker genes that can be used for transformation of most filamentous fungi and yeasts such as acetamidase genes or cDNAs (the *amdS* genes or cDNAs from *A.nidulans*, *A.oryzae*, or *A.niger*), or genes providing resistance to antibiotics like G418, hygromycin, phleomycin or benomyl resistance (benA). Alternatively, more specific selection markers can be used such as auxotrophic markers which require corresponding mutant host strains: e.g. *URA3* (from *S.cerevisiae* or analogous genes from other yeasts), *pyrG* (from *A.nidulans* or *A.niger*) or *argB* (from *A.nidulans* or *A.niger*). In a more preferred embodiment, the selection marker is deleted from the transformed host cell after introduction of the expression construct in accordance with the methods described in EP-A-0 635 574, so as to obtain transformed host cells capable of producing the polypeptide which are free of selection marker genes.

Other markers include ATP synthetase, subunit 9 (*oli*C), orotidine-5'-phosphate-decarboxylase (*pvr*A), the bacterial G418 resistance gene (this may also be used in yeast, but not in fungi), the ampicillin resistance gene (*E. coli*), the neomycin resistance gene (*Bacillus*) and the *E. coli uid*A gene, coding for β-glucuronidase (GUS). Vectors may be used *in vitro*, for example for the production of RNA or used to transfect or transform a host cell.

For most filamentous fungi and yeast, the expression construct is preferably integrated in the genome of the host cell in order to obtain stable transformants. However, for certain yeasts also suitable episomal vector systems are available into which the expression construct can be incorporated for stable and high level expression, examples thereof include vectors derived from the 2μ and pKD1 plasmids of *Saccharomyces* and

Kluyveromyces, respectively. In case the expression constructs are integrated in the host cells genome, the constructs are either integrated at random loci in the genome, or at predetermined target loci using homologous recombination, in which case the target loci preferably comprise a highly expressed gene. A highly expressed gene is herein defined as a gene whose mRNA can make up at least 0.05% (w/w) of the total cellular mRNA, e.g. under induced conditions, or alternatively, a gene whose gene product can make up at least 1% (w/w) of the total cellular protein, or, in case of a secreted gene product, can be secreted to a level of at least 0.1 g/l. A number of examples of suitable highly expressed genes is provided herein below.

An expression construct for a given host cell will usually contain the following elements operably linked to each other in a consecutive order from the 5'-end to 3'-end relative to the coding strand of the sequence encoding the polypeptide of the first aspect: (1) a promoter sequence capable of directing transcription of the DNA sequence encoding the polypeptide in the given host cell, (2) optionally, a signal sequence capable of directing secretion of the polypeptide from the given host cell into the culture medium, (3) the DNA sequence encoding a mature and preferably active form of the polypeptide, and preferably also (4) a transcription termination region (terminator) capable of terminating transcription downstream of the DNA sequence encoding the polypeptide.

Enhanced expression of the polynucleotide encoding the polypeptide of the invention may also be achieved by the selection of heterologous regulatory regions, e.g. promoter, secretion leader and terminator regions, which serve to increase expression and, if desired, secretion levels of the protein of interest from the chosen expression host and/or to provide for the inducible control of the expression of the polypeptide of the invention.

Aside from the promoter native to the gene encoding the polypeptide of the invention, other promoters may be used to direct expression of the polypeptide of the invention. The promoter may be selected for its efficiency in directing the expression of the polypeptide of the invention in the desired expression host.

A variety of promoters[3,4] can be used that are capable of directing transcription in the host cells of the invention. Preferably the promoter sequence is derived from a highly expressed gene as previously defined. Examples of preferred highly expressed genes from which promoters are preferably derived and/or which are comprised in preferred predetermined target loci for integration of expression constructs, include but are not

limited to genes encoding glycolytic enzymes such as triose-phosphate isomerases (TPI), glyceraldehyde-phosphate dehydrogenases (GAPDH), phosphoglycerate kinases (PGK), pyruvate kinases (PYK), alcohol dehydrogenases (ADH), as well as genes encoding amylases, glucoamylases, xylanases, cellobiohydrolases, ß-galactosidases, alcohol (methanol) oxidases, elongation factors and ribosomal proteins. Specific examples of suitable highly expressed genes include e.g. the *LAC4* gene from *Kluyveromyces* sp., the methanol oxidase genes (*AOX* and *MOX*) from *Hansenula* and *Pichia*, respectively, the glucoamylase (*glaA*) genes from *A.niger* and *A.awamori*, the *A.oryzae* TAKA-amylase gene, the *A.nidulans gpdA* gene and the *T.reesei* cellobiohydrolase genes.

Examples of strong constitutive and/or inducible promoters which are preferred for use in fungal expression hosts are those which are obtainable from the fungal genes for xylanase (*xln*A), phytase, ATP-synthetase, subunit 9 (*oli*C), triose phosphate isomerase (*tpi*), alcohol dehydrogenase (*Adh*A), α-amylase (*amy*), amyloglucosidase (AG - from the *gla*A gene), acetamidase (*amd*S) and glyceraldehyde-3-phosphate dehydrogenase (*gpd*) promoters.

Examples of strong yeast promoters are those obtainable from the genes for alcohol dehydrogenase, lactase, 3-phosphoglycerate kinase and triosephosphate isomerase.

Examples of strong bacterial promoters are the α-amylase and *SP02* promoters as well as promoters from extracellular protease genes.

Host cells and Expression

Preferably the polypeptide is produced as a secreted protein in which case the DNA sequence encoding a mature form of the polypeptide in the expression construct is operably linked to a DNA sequence encoding a signal sequence. Preferably the signal sequence is native (homologous) to the DNA sequence encoding the polypeptide. Alternatively the signal sequence is foreign (heterologous) to the DNA sequence encoding the polypeptide, in which case the signal sequence is preferably endogenous to the host cell in which the DNA sequence is expressed. Examples of suitable signal sequences for yeast host cells are the signal sequences derived from yeast α-factor genes. Similarly, a suitable signal sequence for filamentous fungal host cells is e.g. a signal sequence derived from a filamentous fungal (gluco)amylase gene, e.g. the *A.niger glaA* gene. This may be used in combination with the amyloglucosidase (AG) promoter itself, as well as in combination

with other promoters. Hybrid signal sequences may also be used with the context of the present invention.

Preferred heterologous secretion leader sequences are those originating from the fungal amyloglucosidase (AG) gene (*glaA* - both 18 and 24 amino acid versions e.g. from *Aspergillus*), the α-factor gene (yeasts e.g. *Saccharomyces* and *Kluyveromyces*) or the α-amylase gene (*Bacillus*).

Downstream of the DNA sequence encoding the polypeptide, the expression construct preferably contains a 3' untranslated region containing one or more transcription termination sites, also referred to as a terminator. The origin of the terminator is less critical. The terminator can e.g. be native to the DNA sequence encoding the polypeptide. However, preferably a yeast terminator is used in yeast host cells and a filamentous fungal terminator is used in filamentous fungal host cells. More preferably, the terminator is endogenus to the host cell in which the DNA sequence encoding the polypeptide is expressed.

In a further aspect the invention provides a process for preparing polypeptides according to the invention which comprises cultivating a host cell transformed or transfected with an expression vector as described above under conditions to provide for expression by the vector of a coding sequence encoding the polypeptides, and recovering the expressed polypeptides.

A further aspect of the invention thus provides host cells transformed or transfected with or comprising a polynucleotide or vector of the invention. Preferably the polynucleotide is carried in a vector for the replication and expression of the polynucleotide. The cells will be chosen to be compatible with the said vector and may for example be prokaryotic (for example bacterial), fungal, yeast or plant cells.

Depending on the nature of the polynucleotide encoding the polypeptide of the invention, and/or the desirability for further processing of the expressed protein, eukaryotic hosts such as yeasts or fungi may be preferred. In general, yeast cells are preferred over fungal cells because they are easier to manipulate. However, some proteins are either poorly secreted from yeasts, or in some cases are not processed properly (e.g. hyperglycosylation in yeast). In these instances, a fungal host organism should be selected.

A heterologous host may also be chosen wherein the polypeptide of the invention is produced in a form which is substantially free from other pectin-degrading enzymes. This may be achieved by choosing a host which does not normally produce such enzymes such

as *Kluyveromyces lactis*.

The invention encompasses processes for the production of the polypeptide of the invention by means of recombinant expression of a DNA sequence encoding the polypeptide. For this purpose the DNA sequence of the invention can be used for gene amplification and/or exchange of expression signals, such as promoters, secretion signal sequences, in order to allow economic production of the polypeptide in a suitable homologous or heterologous host cell. A homologous host cell is herein defined as a host cell which is of the same species or which is a variant within the same species as the species from which the DNA sequence is derived.

Suitable host cells are preferably prokaryotic microorganisms such as bacteria, or more preferably eukaryotic organisms, for example fungi, such as yeasts or filamentous fungi, or plant cells.

Bacteria from the genus *Bacillus* are very suitable as heterologous hosts because of their capability to secrete proteins into the culture medium. Other bacteria suitable as hosts are those from the genera *Streptomyces* and *Pseudomonas*. A preferred yeast host cell for the expression of the DNA sequence encoding the polypeptide is of the genera *Saccharomyces*, *Kluyveromyces*, *Hansenula*, *Pichia*, *Yarrowia*, and *Schizosaccharomyces*. More preferably a yeast host cell is selected from the group consisting of the species *Saccharomyces cerevisiae*, *Kluyveromyces lactis* (also known as *Kluyveromyces marxianus* var. *lactis*), *Hansenula polymorpha*, *Pichia pastoris*, *Yarrowia lipolytica*,and *Schizosaccharomyces pombe*.

Most preferred for the expression of the DNA sequence encoding the polypeptide are, however, filamentous fungal host cells. Preferred filamentous fungal host cells are selected from the group consisting of the genera *Aspergillus*, *Trichoderma*, *Fusarium*, *Penicillium*, *Acremonium*, *Neurospora*, *Thermoascus*, *Myceliophtora*, *Sporotrichum*, *Thielavia*, and *Talaromyces*. More preferably a filamentous fungal host cell is of the species *Aspergillus oyzae*, *Aspergillus sojae*, *Aspergillus nidulans*, species from the *Aspergillus niger* Group (as defined by Raper and Fennell, The Genus *Aspergillus*, The Williams & Wilkins Company, Baltimore, pp 293-344, 1965). These include but are not limited to *Aspergillus niger*, *Aspergillus awamori*, *Aspergillus tubigensis*, *Aspergillus aculeatus*, *Aspergillus foetidus*, *Aspergillus nidulans*, *Aspergillus japonicus*, *Aspergillus oryzae* and *Aspergillus ficuum*, and further consisting of the species *Trichoderma reesei*, *Fusarium graminearum*,

Penicillium chrysogenum, *Acremonium alabamense*, *Neurospora crassa*, *Myceliophtora thermophilum*, *Sporotrichum cellulophilum*, and *Thielavia terrestris*.

Examples of preferred expression hosts within the scope of the present invention are fungi such as *Aspergillus* species (described in EP-A-184,438 and EP-A-284,603) and *Trichoderma* species; bacteria such as *Bacillus* species (described in EP-A-134,048 and EP-A-253,455), e.g. *Bacillus subtilis, Bacillus licheniformis, Bacillus amyloliquefaciens, Pseudomonas* species; and yeasts such as *Kluyveromyces* species (described in EP-A-096,430 e.g. *Kluyveronmyces lactic* and EP-A-301,670) and *Saccharomyces* species, e.g. *Saccharomyces cerevisiae.*

Culture of host cells and Recombinant production

According to the present invention, the production of the polypeptide of the invention can be effected by the culturing of microbial expression hosts, which have been transformed with one or more polynucleotides of the present invention, in a conventional nutrient fermentation medium.

The recombinant host cells according to the invention may be cultured using procedures known in the art. For each combination of a promoter and a host cell, culture condition are available which are conducive to the expression the DNA sequence encoding the polypeptide. After reaching the desired cell density or titre of the polypeptide the culture is stopped and the polypeptide is recovered using known procedures.

The fermentation medium can comprise a known culture medium containing a carbon source (e.g. glucose, maltose, molasses, etc.), a nitrogen source (e.g. ammonium sulphate, ammonium nitrate, ammonium chloride, etc.), an organic nitrogen source (e.g. yeast extract, malt extract, peptone, etc.) and inorganic nutrient sources (e.g. phosphate, magnesium, potassium, zinc, iron, etc.). Optionally, an inducer (e.g. apple MHR, pectin or xylogalacturonan) may be included.

The selection of the appropriate medium may be based on the choice of expression host and/or based on the regulatory requirements of the expression construct. Such media are well-known to those skilled in the art. The medium may, if desired, contain additional components favouring the transformed expression hosts over other potentially contaminating microorganisms.

The fermentation can be performed over a period of 0.5-20 days in a batch, continuous

or fed-batch process suitably at a temperature in the range of between 0 and 45°C and, for example, a pH between 2 and 10. Preferred fermentation conditions are a temperature in the range of between 20 and 37°C and/or a pH between 3 and 9. The appropriate conditions are usually selected based on the choice of the expression host and the protein to be expressed.

After fermentation, the cells can be removed from the fermentation broth by means of centrifugation or filtration. After removal of the cells, the polypeptide of the invention may then be recovered and, if desired, purified and isolated by conventional means.

D. Methods of Processing Plant or Pectin-containing Materials

Plant and pectin-containing materials include plant pulp, parts of plants and plant extracts. In the context of this invention an extract from a plant material is any substance which can be derived from plant material by extraction (mechanical and/or chemical), processing or by other separation techniques. The extract may be juice, nectar, base, or concentrates made thereof. The plant material may comprise or be derived from vegetables, e.g., carrots, celery, onions, legumes or leguminous plants (soy, soybean, peas) or fruit, e.g., pome or seed fruit (apples, pears, quince etc.), grapes, tomatoes, citrus (orange, lemon, lime, mandarin), melons, prunes, cherries, black currants, redcurrants, raspberries, strawberries, cranberries, pineapple and other tropical fruits, trees and parts thereof (e.g. pollen, from pine trees). According to this invention, apples and apple juice are especially preferred.

The polypeptides of the invention can thus be used to treat plant material including plant pulp and plant extracts. For example, they may be used to treat apple pulp and/or raw juice during the production of apple juice. They may also be used to treat liquid or solid foodstuffs or edible foodstuff ingredients. Conveniently the polypeptide of the invention is combined with suitable (solid or liquid) carriers or diluents including buffers to produce a composition or enzyme preparation.

The polypeptide is typically stably formulated either in liquid or dry form. Typically, the product is made as a composition which will optionally include, for example, a stabilising buffer and/or preservative. The compositions may also include other enzymes capable of digesting plant material or pectin, for example other pectinases such as an endo-arabinanase, rhamnogalacturonases, and/or polygalacturonase. For certain

applications, immobilization of the enzyme on a solid matrix or incorporation on or into solid carrier particles may be preferred. The composition may also include a variety of other plant material-degrading enzymes, for example cellulases and other pectinases.

The polypeptides and compositions of the invention may therefore be used in a method of processing plant material to degrade or modify the pectin constituents of the cell walls of the plant material[2].

Typically, the polypeptides of the invention are used as a composition/ enzyme preparation as described above. The composition will generally be added to plant pulp obtainable by, for example mechanical processing such as crushing or milling plant material. Incubation of the composition with the plant will typically be carried out for at time of from 10 minutes to 5 hours, such as 30 minutes to 2 hours, preferably for about 1 hour. The processing temperature is preferably 10-55°C, e.g. from 15 to 25°C, optimally about 20°C and one can use 10-300g, preferably 30-70g, optimally about 50g of enzyme per ton of material to be treated. All the enzyme(s) or their compositions used may be added sequentially or at the same time to the plant pulp. Depending on the composition of the enzyme preparation the plant material may first be macerated (e.g. to a purée) or liquefied. Using the polypeptides of the invention processing parameters such as the yield of the extraction, viscosity of the extract and/or quality of the extract can be improved.

Alternatively, or in addition to the above, a polypeptide of the invention may be added to the raw juice obtained from pressing or liquefying the plant pulp. Treatment of the raw juice will be carried out in a similar manner to the plant pulp in respect of dosage, temperature and holding time. Again, other enzymes such as those discussed previously may be included. Typical incubation conditions are as described in the previous paragraph. Once the raw juice has been incubated with the polypeptides of the invention, the juice is then centrifuged or (ultra) filtered to produce the final product.

A composition containing a polypeptide of the invention may also be used during the preparation of fruit or vegetable purees.

The end product of these processes is typically heat-treated at 85°C for a time of from 1 minute to 1 hour, under conditions to partially or fully inactivate the polypeptides of the invention.

Due to the highly specific action on pectins the polypeptides of the invention may also be used to prepare pectins with modified characteristics, e.g. modified gelation capacities

for specific applications.

The polypeptides of the invention may also be added to animal feeds rich in pectin or xylogalacturonan, e.g. soy-containing food, to improve the breakdown of the plant cell wall leading to improved utilisation of the plant nutrients by the animal. The polypeptides of the invention may be added to the feed or silage if pre-soaking or wet diets are preferred. Advantageously, the polypeptides of the invention may continue to degrade xylogalacturonans in the feed *in vivo*. Fungal derived polypeptides of the invention in particular generally have lower pH optima and are capable of releasing important nutrients in such acidic environments as the stomach of an animal. The invention thus also contemplates (e.g. animal) feeds or foodstuffs comprising one or more polypeptides of the invention.

The polypeptides of the invention may also be used during the production of milk substitutes (or replacers) from soy bean. These milk substitutes can be consumed by both humans and animals. A typical problem during the preparation of these milk substitutes is the high viscosity of the soy bean slurry, resulting in the need for an undesirable dilution of the slurry to a concentration of dry solids of 10 to 15%. An enzyme preparation containing a polypeptide of the invention can be added to, or during the processing of, the slurry, enabling processing at a higher concentration (typically 40 to 50%) dry solids. The enzyme may also be used in the preparation of savoury product(s), e.g. from soy bean.

Assays for Pectin Degrading Enzymes

The novel assays and substrates described herein have allowed identification and confirmation of endo-xylogalacturonase activity. However, these assays can be used to detect other pectin degrading enzymes, whether or not they have endo-xylogalacturonase activity.

The substrate that can be used for this assay can comprise gum tragacanth which has been treated with a strong acid. A preferred acid is trifluoroactetic acid (TFA). The gum tragacanth may be optionally saponified and/or it may have been treated with an alkali, for example an alkali metal hydroxide, for example NaOH.

Another aspect of the invention relates to an assay for identifying or detecting a polypeptide which is able to degrade pectin. The activity may be an endo-xylogalacturonase or, may be pectin lyase, polygalacturonase, esterase, cellulase,

xyloglucanase, galactonase, arabinanase or rahamnogalacturonase. The assay may comprise:

a. providing, as a substrate for a candidate compound (usually a polypeptide) the substrate described in the previous paragraph; and

b. contacting the substrate with the candidate compound, and detecting whether any reducing carbohydrates are produced.

The amount of these reducing carbohydrates can be measured. If necessary, they can then be compared to the amount of the carbohydrates produced in a control experiment, in the absence of candidate compound.

The measurement may involve a BCA assay. This may comprise measuring the amount of Cu(II) reduced to Cu(I) by the reducing carbohydrates present. This may be by contact with bicinchoninic acid (BCA), and determining the amount of BCA-Cu(I) complex formed.

The invention will now be described with reference to the following Examples which are intended to be illustrative only and not limiting. In the Figures which accompany the Examples:

Figure 1 is a diagram of the hypothetical structure of the prevailing population of apple MHR (modified hairy region) having the highest molecular weight (subunit I is xylogalacturonan, subunit II is the backbone rich in arabinan side chains, subunit III is rhamnogalacturonase oligomers. The distribution of acetyl groups is not presented but major parts are thought to be located within subunit III. Key: GalA = galacturonic acid; rham = rhamnose; gal=galactose; xyl = xylose; ara = arabinose);

Figure 2 is a map of the vector pCVlacK according to the invention (construction described in Example 1);

Figure 3 is a graph illustrating an HPAEC of xylogalacturonan after degradation by xylogalacturonase (a polypeptide of the invention);

Figure 4 is a graph illustrating an HPSEC of xylogalacturonan before and after degradation by a xylogalacturonase;

Figure 5 is a graph illustrating a Maldi-ToF mass spectrum of the products of complete degradation of xylogalacturonan by a xylogalacturonase;

Figures 6A-G are graphs of HPSEC elutions showing degradation of MHR-S by endo-arabinanase, rhamnogalacturonase and xylogalacturonase, separately and in

combination;

Figures 7 and 8 are graphs of a HPSEC and HPAEC, respectively, elution patterns showing degradation of soy pectin by xylogalacturonase; and

Figure 9 is a diagram showing the multiple alignment of the part of the PG, XghA and RHG sequences (amino acids identical to the XghA sequence are replaced by a dot) and an introduction of gaps to obtain an optimal alignment are indicated (-). Conserved amino acids in all plant, fungal and prokaryotic PG's are shaded. Key: Atub = *A. tubigensis*; Anig = *A. niger*; Aac = *A. aculeatus*.

EXAMPLE 1

Construction of an *Aspergillus tubigensis* cDNA expression library

Example 1.1: Construction of an expression vector

Starting vector pGBHSA20 (CBS 997.96) contains the promoter and terminator sequence of the lactase gene (*lac4*) of *K. lactis*, a G418 selection marker and the *E. coli* plasmid pTZ18r for propagation in this host. The *K. lactis* KARSCEN cassette[17] (a gift from Dr. A.A. Winkler, Dept. of Cell Biology and Genetics, Leiden University, The Netherlands) was cloned in a unique *Sma*I site of this vector. The resulting vector was named pCVlacK (Figure 2). The unique *Hin*dIII and *Xho*I sites flanking the *lac4* promoter and terminator, respectively, can be used as cloning sites for cDNA synthesized from *Aspergillus tubigensis* poly(A) RNA.

Example 1.2: Isolation of poly(A) RNA and cDNA synthesis

Aspergillus tubigensis conidia were inoculated in triplicate at a density of 10^6 spores/ml in 300 ml of medium containing (per liter): 6 g $NaNO_3$, 0.5 g KCl, 1.5 g KH_2PO_4, 0.5 g $MgSO_4$ (pH6.5), 1 ml 1000x Timberlake spore elements (per ml, 50mg EDTA, 22mg $FeSO_4.7H_2O$, 5 mg $MnCl_2.2H_2O$, 22mg $ZnSO_4.7H_2O$, 1.6mg $CuSO_4.5H_2O$, 1.7mg $CoCl_2.6H_2O$, 1.5mg $Na_2MoO_4.2H_2O$, 11mg H_3BO_3, adjusted to pH 6.5) and 10ml 100x Timberlake vitamins (per ml, 0.2mg thiamine-HCl, 0.2mg riboflavin 0.2mg nicotinamide, 1mg pyridoxine-HCl, 0.02mg pantothenic acid, 0.4µm biotin, adjusted to pH 5 to 6), 1 g yeast extract, 5 g Soyoptim™ (defatted, toasted soy bean meal from Societé Industrièlle Oleagineux, France). The cultures were incubated in a rotary shaker at 28°C,

150 rpm. The mycelium of one culture was harvested at 10 hours after inoculation, mycelium of the other two cultures at 16 and 24 hours after inoculation. From 1 g rinsed and squeezed mycelium total RNA was isolated by the RNAzol method (Cinna/Biotecx). Poly(A) RNA was isolated using Qiagen™ oligotex columns (Westburg). Equal amounts of poly(A) RNA at time-points of 10, 16 and 24 hours were pooled. cDNA was synthesized using the ZAP-cDNA synthesis kit (Stratagene™) with the following modifications: the first-strand synthesis was done with Superscript II reverse transcriptase (GibcoBRL). To 7.5 µg poly(A) RNA, 2 µl linker-primer and RNAse-free water was added to a final volume of 28.5 µl. This mixture was incubated for 10 minutes at 70°C and chilled on ice. The following components were added: 10 µl 5x first strand buffer, 5 µl 0.1 M DTT, 3 µl first-strand methyl nucleotide mixture and 1 µl RNAse block. This was incubated for 10 minutes at 25°C, followed by 2 minutes at 42°C. Subsequently, 2.5 µl Superscript™ II RT (200 U/µl) was added, mixed and incubated for 50 minutes at 42°C. A second modification of the protocol was ligation of a *Hin*dIII adaptor instead of the *Eco*RI adaptor.

The cDNA pool was size separated using a Sephacryl S-500 column. The first fraction eluted from the column did not contain any cDNA but the second and third fraction contained the largest sized cDNA. Subsequent fractions were supposed to contain relatively higher amounts of non-full length cDNA and were of no use for construction of the library. The cDNA of fractions 2 and 3 was ligated into the *Hin*dIII and *Xho*I sites of expression vector pCVlacK (see Figure 2) using the Clontech Ligation Express™ kit. Each ligation mixture was transformed in two batches to electrocompetent *E. coli* XL-Blue MRF'cells. The four transformation suspensions were plated onto 32 agar plates (LB + 50µg/ml ampicillin). After 16 hours of incubation at 37°C, 7366 transformants were obtained. Bacteria were collected by pouring 2.5 ml LB medium onto a plate and then scraping off the cells; 0.5 ml of cell suspension was added to glycerol and stored at -80°C; the remaining 2 ml was used for DNA isolation (Qiagen™ Spin miniprep kit). In case a low number of transformants per plate was found, the 2.5 ml was transferred to a second, third or fourth plate. This yielded 22 pools of about 325 individual transformants. Equal amounts of DNA of each pool were combined for use in *K. lactis* transformation.

Example 1.3: Transformation of the expression library into *K. lactis*

An overnight culture of *K. lactis* strain CBS 2359 grown in YPD (10 g/l yeast extract, 20 g/l Bacto-peptone, 20 g/l glucose) at 30°C was diluted 3000-, 600-, 300- and 100-fold in 150 ml of fresh YPD and incubated for 6 hours at 30°C, 160 rpm in a rotary shaker. The culture with an optical density of 0.7-1.0 was used to prepare electrocompetent cells[6].

Electrocompetent cells were transformed with 1 µg pooled DNA of the *E. coli* library. Electroporation was performed using a Biorad Genepulser™ with settings at 1.4 kV, 200 Ohm and 25 µF. Transformants were selected on double layer YPD plates (YPD with 20 g/l Bacto-agar): the bottom layer contained 50 µg/ml G418, the top layer was non-selective. 660 µL of transformation mix was plated onto 80 double layer plates. Aliquots of 1.5 and 15 µL were pipetted onto the plates. About 10,000 transformants were obtained.

EXAMPLE 2

Substrate preparation

Example 2.1: Preparation of MHR-S from apples

Modified hairy regions (MHR) from apples were isolated as a filter retentate after treatment of apples with pectinase, and subsequently the MHR was saponified resulting in MHR-S[8].

Example 2.2: Synthesis of xylogalacturonan from gum tragacanth

5 g of gum tragacanth (Sigma, St. Louis, MO, USA) was suspended in 990 mL of ice-cold distilled water. To this solution 10 mL of an ice-cold 5 M NaOH solution was added. After 24 h, the saponified gum tragacanth (sGT) was dialysed extensively against distilled water at 4°C , and concentrated under reduced pressure to 1 L. For mild acid hydrolysis, 7.65 mL of trifluoric acid (TFA) was added to this sGT solution. Final concentration of TFA in the solution was 0.1 M. The sGT/TFA solution was heated to boiling point in a microwave and subsequently incubated in a boiling water bath for 1 h. Finally the hydrolysate was dialyzed extensively against distilled water at 4°C and freeze-dried. This procedure yielded 2.61 g of material (further referred to as xylogalacturonan).

Example 2.3: Characterization of xylogalacturonan

To determine the sugar composition both from the original gum tragacanth as well as of the xylogalacturonan, samples were hydrolyzed with 1 M H_2SO_4 (100°C, 3 h)[8] and neutral sugars were converted to their alditol acetates in order to quantify the individual sugars by Gas Chromatography (GC). The uronic acid content of the hydrolysate was determined colorimetrically using *m*-hydroxybiphenyl[8]. The sugar composition in mol% of the original gum tragacanth (GT) in comparison with the xylogalacturonan (XG) is shown in Table 1 below.

Table 1

	Sugar (mol%)						
Substrate	Rhamnose (Rha)	Fucose (Fuc)	Arabinose (Ara)	Xylose (Xyl)	Galactose (Gal)	Glucose (Glc)	Galacturonan (GalA)
GT	2	5	26	17	7	7	36
XG	3	0	1	25	9	6	56

The mild acid hydrolysis removed effectively arabinosyl and fucosyl residues from this polysaccharide, whereas the GalA:Xyl ratio was more or less unaltered.

The degree of acetyl and methyl esterification of gum tragacanth was estimated by High Pressure Liquid Chromatography (HPLC)[7]. The degree of methylation and acetylation of gum tragacanth is approximately 75% and 20%, respectively (calculated as mol methyl or actyl groups per mol of GalA). Saponification of the gum removed all methyl and acetyl groups. Molecular weight distribution of the polymers was performed by High Pressure Size Exclusion Chromatography (HPSEC) on three Bio-Gel TSK columns (40XL, 30XL, and 20XL) in series[8], Mild acid hydrolysis of saponified gum tragacanth was accompanied with a decrease in molecular weight, although the resulting material still had a high molecular weight. Based on HPSEC elution profiles, xylogalacturonan has an estimated molecular mass of approximately 1,100 kDa, using pullulan reference compounds.

Xylogalacturonan effectively proved to be resistant to enzymatic degradation by all tested endo-polygalacturonases and rhamnogalacturonases.

EXAMPLE 3

Screening the library with a BCA assay

Example 3.1: Growth of the transformants and enzyme preparation

From the about 10 000 *K. lactis* transformants produced in Example 1.3, 3,500 individual colonies were picked and transferred to separate wells of multi-well plates. The transformants in the multiwell plates were grown for 48 hours at 30°C in medium I (per 500 mL of H_2O (pH 6.0): mannitol, 10.00 g; $NH_4H_2PO_4$, 1.50 g; KH_2PO_4, 0.25 g; $(NH_4)_2SO_4$, 0.50 g; $CaCl_2.2H_2O$, 0.01 g; $MgSO_4.7H_2O$, 0.15 g: trace elements H_3BO_3, 375 µg; $CuSO_4.5H_2O$, 40 µg; KI, 75 µg; $MnSO_4.4H_2O$, 300 µg; Na_2MoO_4, 150 µg; $ZnSO_4.7H_2O$, 300 µg; $FeCl_3.6H_2O$, 200 µg) and vitamins Ca-pantothenate, 500 µg; thiamine, 500 µg; myo-inositol, 500 µg; pyridoxine, 500 µg; nicotinic acid, 500 µg; biotin, 5 µg) containing 80 ng/mL of the antibiotic G418 and the 35 plates were stored as 15% glycerol stocks.

These transformants were used to inoculate a new set of 35 multi-well plates containing 200 µL of the same medium with 80 ng/mL G418 with a replica plater. The *K. lactis* transformants were grown for two days at 30°C in a stove. The cells were precipitated by centrifugation at 3000 rpm in a Hermle™ zk380 centrifuge.

Example 3.2: Substrate degradation

Carefully, 25 µL of supernatant of each well was pipetted to a new multiwell plate, and 25 µL of a 0.2% solution of substrate (either MHR-S or xylogalacturonan or sGT/TFA from Example 2.2) or in 100 mM NaOAc buffer pH 5.0 was added. After incubation overnight in a stove at 30°C, the increase of reducing carbohydrates was measured with the BCA assay.

Example 3.3: The BCA assay

The BCA assay is based on the reduction of Cu(II) to Cu(I) by reducing carbohydrate mono- and oligomers. A complex is formed of bicinchoninic acid (BCA) and Cu(I). This complex produces an intense purple colour, which can be measured spectrophotometrically. This colour increases with an increasing reducing carbohydrate

concentration. The method used in this invention is a modification of a known method[9] but was used for screening purposes.

The procedure consisted of mixing 10 µL of reducing carbohydrate containing sample from Example 3.2 with 90 µL of water and 100 µL of BCA reagent together in a multi-well plate. BCA reagent was made freshly each day by mixing two solutions, A and B, 1:1 (v/v) together. Solution A consisted of 54.28 g Na_2CO_3, 24.20 g $NaHCO_3$ and 1.942 g Na_2BCA per liter of destilled water. Solution B consisted of 1.248 g $CuSO_4.5H_2O$ and 1.262 g *L*-serine per liter of distilled water. The plate containing sample, reagent and water was incubated for one hour in a stove at 80°C with a lid on the plate. After cooling the plate for 15 minutes, the absorbance at 550 nm was measured using a multi-well plate reader (SLT lab instruments, Austria; model EAR 400). Testlines with galactose showed that the assay was linear in the range from 0 to 125 µM galactose.

Transformants that produced in the BCA assay an absorbance 0.1 unit higher than the blank were checked for xylogalacturonan-degrading capabilities by growing them again and repeating the BCA assay using xylogalacturonan as a substrate. Three xylogalacturonase producing transformants were found.

EXAMPLE 4

Example 4.1: Characterization of xylogalacturonase encoding cDNA

All plasmid inserts of these three transformants were identical, as was found after analysis of the restriction patterns of these inserts. *K. lactis* transformant 27E8, exhibiting xylogalacturonase activity, was used to isolate the pCVlacK expression plasmid by a glass beads method[10]. After transformation and propagation of this plasmid in *E. coli*, the cDNA insert was excised from pCVlacK with a *Hin*dIII/*Xho*I digestion. This digestion released a 1.0 and 0.4 kb fragment, due to an internal *Hin*dIII site as appeared from the nucleotide sequence later on. The DNA sequence of the cDNA insert was determined on both strands using 5'- and 3'-specific primers to the *lac4* regulating sequences, and primers based on the cDNA sequence. The DNA sequence of the cDNA insert is presented in SEQ ID No. 1, together with the deduced amino acid sequence. Upstream of the ATG translation start codon, 20 nucleotides of 5'-untranslated sequence are present. Downstream the TAA stop codon 130 nucleotides non-translated sequence followed by the poly-A tail were found. The open reading frame of 1218 nucleotides (*xghA*) encodes a protein of 406 amino acids,

presented in SEQ ID No. 2, named XghA. The potential cleavage site of the signal sequence, predicted according to the (-3,-1) rule[11], is between position 18 and 19. The ORF, initiating with an ATG codon and terminating with a TAA codon is thus preceded by a 20 base pair 5'-noncoding region and followed by a 130 base pair 3'-noncoding region and a poly(A) tail. The sequence TCATCATGGC covering the ATG start codon closely resembles the contents of sequence for initiation of translation in higher eukaryotes[14]. The *xghA* cDNA encodes an apparent signal sequence of 18 amino acids at the amino terminus, with a signal peptidase cleavage site between Ala^{18} and Ala^{19}. Two potential N-glycosylation sites have been found at Asn^{178}-Ser-Thr and Asn^{301}-Val-Thr.

Comparison of the amino acid sequence to protein databases showed homology to polygalacturonase sequences of prokaryotes, fungi, plants and to rhamnogalacturonases A and B of *Aspergillus*. A comparison of the XghA amino acid sequence has been made using sequences from the EMBL data library. XghA showed 31 to 39% similarity to the endo-PG's and 44% similarity to the exo-PG of *A. tubigensis*. Similarity of XghA to 2RHG-A's was 30% (*A.niger*) and 32% (*A. aculeatus*) whereas similarity to RGH-B of *A. niger* was only very limited.

Figure 9 shows the analysis of multiple alignments of Xgh to the PG's and to the RHG-A's. The multiple alignment shows four domains of conserved amino acids, which were first described for polygalacturonases of plant, fungal and bacterial origin[15]. When all the PG's that were recovered from a database search were aligned, only four small stretches of amino acids are conserved: NXD, DD, HG and RXK (shaded in Figure 9, where X represents a variable amino acid). Essential amino acids thought to be involved in the hydrolysis reaction are one of the three aspartic acid residues of domain I and II and the histidine of domain III. These domains are fully conserved in XghA. It is postulated that the fourth domain contains amino acids that are involved in substrate binding. The arginine residue of this domain is a glycine residue in XghA. The domains are less conserved in the RHG sequences, as only two of the three aspartic acid residues are conserved and the histidine is replaced by glycine.

Example 4.2: Southern blot analysis

The copy number of the *XghA* gene was determined by southern blot analysis of genomic DNA of *A. tubigensis* digested with several enzymes (results not shown).

Hybridization under stringent (65°C and 0.2 x SSC) and less stringent conditions (60°C and 1 x SSC) with a 1.0 kb *Hin* dIII fragment of *xghA* clearly showed single hybridizing fragments. This demonstrates that the *xghA* gene is present as a single copy in the *A. tubigensis* genome.

EXAMPLE 5

Expression of the enzyme

K. lactis transformants expressing the cDNA of endo-xylogalacturonase were transferred from multiwell plate glycerol stocks to reagent tubes: 10 µL glycerol stock was added to 1-2 mL of medium I (see Example 3.1) with 80 ng/mL of G418. These cultures were grown at 30°C in a rotary incubator at 200 rpm for two days and used to inoculate Erlenmeyer flasks containing 20 mL of this same medium supplemented with antibiotic. For larger scale production of the enzyme, these cultures were used to inoculate 500 mL of the same medium supplemented with antibiotic in 1 L Erlenmeyer flasks. Cells were grown at 30°C in a rotary incubator at 200 rpm for two days. Cultures were centrifuged to precipitate cells, the supernatant was used for the purification.

The crude enzyme preparation (350 mL) was preconcentrated on a Hitrap™ Q ion-exchange column (Pharmacia Biotech, Sweden) with a flow rate of 0.3 mL/min. Elution was performed on a FPLC system (Pharmacia Biotech, Sweden) with a salt gradient using a 20 mM piperazine (pH 5.0) starting buffer (buffer A) and a 0.5 M NaCl in 20 mM piperazine (pH 5.0) elution buffer (buffer B). The following gradient was used: to 10% B in 1 minute, to 35% B in 19 minutes, to 100% B in 2 minutes and 100% B for three more minutes. Activity was checked as described in Example 3 and active fractions were pooled. They were diluted three times with 20 mM piperazine buffer (pH 5.0) and applied on a MiniQ column (Pharmacia Biotech, Sweden). Elution was performed on a Smart system (Pharmacia Biotech, Sweden) with a linear pH gradient from 20 mM piperazine (pH 5.0) starting buffer to 10 mM HCl a flow rate of 0.4 mL/min. Active fractions were pooled and investigated using SDS-PAGE. Upon silver staining of the gel one protein band with a molecular mass of approximately 60 kDa was found. The difference with the predicted MW of about 45 kDa, based on the DNA sequence (see Example 4) is thought to be due to protein glycosylation

EXAMPLE 6

Influences of pH and temperature on enzyme activity

Purified enzyme, obtained as described in Example 5, was used for the characterisation of the enzyme. Measurements at t=0 were used as blanks.

For the determination of the pH stability, the purified enzyme was preincubated without substrate for one hour at a pH range from 2.5 to 8 in McIlvaine buffers. Afterwards the enzyme was incubated with substrate for two hours and the increase in reducing sugars was determined as described in Example 3. The enzyme was stable over a pH range of 3 to 6.

For the determination of the pH and temperature optima, the purified enzyme was incubated with substrate for two hours at a pH range from 2.5 to 8 or a temperature range from 20 to 80°C. After this, the increase in reducing sugars was determined as described in Example 3. The enzyme has an optimum activity at a temperature of 60°C and at a pH of 3.0. The enzyme shows more than 50% of its activity in the pH range of 2.5 to 5.0. The activity at pH 2.5 was still 90% of the maximum value at pH 3.0. Values lower than pH 2.5 were not measured.

EXAMPLE 7

Mode of action of the xylogalacturonase

Degradation of xylogalacturonan (modified gum tragacanth, Example 2.2) by the supernatant of the xylogalacturonase-producing *K. lactis* clone was monitored by high performance anion exchange chromatography (HPAEC) and high performance size exclusion chromatography (HPSEC).

HPAEC was performed using a Dionex carbopack PA1 column of a size of 4 x 250 ml. Elution was performed with 0.1 M NaOH (solution A) and 1 M NaOAc in 0.1 NaOH (solution B). The following gradient was used: from 0 to 62% B in 50 minutes, to 100% B in 5 minutes, followed by 100% B for 5 minutes. The enzyme did not produce xylose (expected at a retention time of 5 minutes) or galacturonic acid (expected at a retention time of 15 minutes), not even after 8 hrs incubation. Only oligomers were released, the smallest oligomer being found at a retention time of about 22 minutes: this was a xylose-galacturonic acid dimer. (In Figure 3, Bottom line (A): t=1h, middle line (B): t=4h, top line (C): t=8h of incubation).

HPSEC was performed using three columns in series: Bio-Gel TSK 40 (300 x 7.5 mm, from Biorad), Bio-Gel TSK 30 XL (300 x 7.5 mm, from Biorad) and TSKGel G 2500 P

XL (300 x 7.8 mm, from TosoHaas). Figure 4 shows that the high molecular weight fraction of xylogalacturonan (left in the picture) was rapidly degraded (the top line (B) represents the polymer before degradation and the bottom line (A) represents the polymer after degradation).

When the degradation products were monitored by Maldi-ToF mass spectrometry the products found were identified and are shown in Table 2.

Table 2

Oligomers	Composition of Degradation Product	Peak Numbers (in Figure 4)	Molecular Weight (Da)
dimer	galAxyl	1	349.1
trimer	galA2xyl	2	525.1
tetramers	galA2xyl2, galA3xyl	3a, 3b	657.3, 701.3
pentamers	galA3xyl2, galA4xyl1	4a, 4b	833.4, 877.4
hexamer	galA4xyl2	5	1009.5
heptamers	galA4xyl3, galA5xyl2	6a, 6b	1141.6, 1185.5
octamer	galA6xyl2	7	1361.5

Two products were thus formed when MHR-S was incubated with XghA. These products appeared even after a short incubation time.

These results show that xylogalacturonan is degraded by the xylogalacturonase in an endo-fashion. The profiles obtained were compared with those from a polygalacturonic acid digest, and it was clear that none of the MHR-S degradation products formed were polygalacturonic acid oligomers. This demonstrates that XghA produces galacturonic acid oligomers substituted with xylose.

Upon incubation of the supernatant of the xylogalacturonase producing *K.lactis* transformant with polygalacturonic acid no degradation of this substrate was observed.

EXAMPLE 8

Complete degradation of MHR-S by xylogalacturonase in combination with other enzymes

To study the degradation of MHR-S 200 µL of a 0.3% MHR-S solution in 50 mM NaOAc buffer pH 5.0, was incubated with 5 µL of the purified xylogalacturonase, 5 µL of endo-arabinanase[12], 5 µL of rhamnogalacturonase[13], or with combinations of these enzymes, added sequentially or at the same time MHR-S without enzyme was used as a control.

The degradation of MHR-S was monitored with HPSEC, as described in Example 7. The results are shown in Figures 6 to 6G where upper line (b) represents the control and the lower line (d) represents the incubations with enzyme Incubations were with:

A: arabinanase;

B: xylogalacturonase;

C: rhamnogalacturonase;

D: endo-arabinanase and xylogalacturonase sequentially;

E: endo-arabinanase and xylogalacturonase combined;

F: endo-arabinanase and rhamnogalacturonase sequentially; and

G: endo-arabinanase rhamnogalacturonase and xylogalacturonase combined.

Figure 6B shows that xylogalacturonase was able to degrade MHR-S: a small shift to lower molecular weight material can be observed. Also the enzymes endo-arabinanase (Figure 6A) and rhamnogalacturonase (Figure 6C) caused some shift in molecular weight. However, somewhat better results are obtained by combining two different enzymes in one incubation (Figures 6D – endo-arabinanase and xylogalacturonase sequentially; 6E - endo-arabinanase and xylogalacturonase combined and 6F -endo-arabinanase and endo-rhamnogalacturonase sequentially). The difference between Figures 6D and 6E is striking: combined addition of endo-arabinanase and xylogalacturonase was much more effective than with sequential addition. Almost complete degradation of the high molecular weight material was possible when the three enzymes were added combined (Figure 6G).

EXAMPLE 9

Improvement of filtration rate by xylogalacturonase in combination with other enzymes

Experiments were done to see if the xylogalacturonase could prevent filterfouling during filtration. Apple MHR-S prepared as described in Example 2 1 was used as a

substrate. A solution of 0.5% in 50 mM acetate buffer pH 4.0 was incubated with a combination of three enzymes: endo-arabinanase, rhamnogalacturonase and xylogalacturonase (ea/rg/xgh), a combination of two enzymes: endo-arabinanase and rhamnoglacturonase (ea/rg), and with xylogalacturonase (xgh) separately for 17 hours at 30°C. The solutions were filtrated using an Amicon device equipped with a 30 kD filter at a pressure of 2 bars. The increase in weight of the filtrate was followed over time. The results are shown in Figure 5.

EXAMPLE 10

Degradation of a soy pectic fraction by xylogalacturonase

An Alkali soluble fraction of soybean meal, 1 MASS, rich in pectic substances was isolated and characterized[16]. A 0.25% solution of 1 MASS in 50mM sodium acetate buffer pH 5, including 0.01% NaN_3 was incubated with xylogalacturonase. The digest obtained after 24 hours of incubation at 30°C was analysed for the molecular weight distribution by HPSEC and the release of oligomeric degradation product by HPAEC. The analyses were performed as described in Example 7.

Figure 7 shows the changes in the molecular weight distribution as measured by HPSEC: in the xylogalacturonase-treated material (curve b) the peak at approximately 20 min, representing the high molecular weight material, decreases to 70% of the value of the starting material (curve a).

In Figure 8 the results of the HPAEC analysis is shown. Comparing the enzyme-treated material (curve b) with the blank (curve a) it can be seen that xylogalacturonase causes the release of the characteristic xylosyl galacturonic acid dimer (marked with an X) and of other unidentified oligomers (peaks to the right side of X), comparable with the peaks appearing in Figure 6 of Example 7.

REFERENCES

1. G Beldman, L.A.M. van den Broeck, H.A. Schols, M.J.F Searle-van Leeuwen, K.M.J van Laere and A.G.J. Voragen (1996) *Biotechnology Letters* 18: 707-712

2. C. Grassin and P. Fauquembergue (1996) *Pectins and Pectinases, Progress in Biotechnology* 14: 453-462

3. Goosen *et al.*, 1992, "Transformation and Gene Manipulation in filamentous fungi: an overview" In: Handbook of Applied Mycology" Vol. 4: "Fungal Biotechnology", D.K.

4. Romanos *et al*, Yeast 8: 423-488 (1992)

5. Arora, R.P. Elander and K.G. Mukerji, eds., Marcel Deker In., New York, pp 151-195

6. K.N. Faber, P. Haima, W. Harder, M. Veenhuis, G. AB (1994), *Current Genetics* 25: 305-310.

7. A.G.J. Voragen, H.A. Schols and W. Pilnik, (1986) *Food Hydrocoll.* 1: 65-70

8. H.A. Schols, M.A. Posthumus, A.G.J. Voragen (1990) *Carbohydrate Research* 206: 117-129

9. J.D. Fox, J.F. Robyt (1991) *Analytical Biochemistry* 195: 93-96

10. M.A. Sobanski and Dickinson (1995) *Biotechnology Techniques* 9: 225-230)

11. G. von Heijne (1986) *Nucleic Acids Research* 14: 4683-4690)

12. G. Beldman, M.J.F. Searle van Leeuwen, G.A. de Ruiter, H.A. Siliha, A.G.J. Voragen (1993) *Carbohydrate Polymers* 20: 159-168

13. H.A. Schols, C.C., J.M. Geraerds, M.J.F. Searle van Leeuwen, F.J.M. Kormelink, A.G.J. Voragen (1990) *Carbohydrate Research* 206:105-115

14. M. Kozak, *et al*, The Scanning Model for Translation: an update, J.Cell.Biol. 108 (1989) 229-241.

15. H.J.D. Bussink *et al*, Curr.Genet. 19(1991), 467-474.

16. Huisman, M.M.H. *et al* (1998) Carbohydr. Polym. 37: 87-95.

17. Winkler A.A. *et al*, PhD thesis, University of Leiden, The Netherlands (1998).

CLAIMS

1. A polypeptide which possesses endo-xylogalacturonase activity.

2 A polypeptide having endo-xylogalacturonase activity which is obtainable from a fungus and possesses endo-xylogalacturonase activity.

3 A polypeptide according to claim 2 wherein said fungus is of the genus *Aspergillus*.

4. A polypeptide according to any preceding claim which comprises the sequence set out in SEQ ID No. 2, or a sequence substantially homologous thereto, or a fragment of either sequence.

5 A polypeptide according to claim 4 wherein the fragment has at least 5 amino acids or the homologous sequence is at least 60% identical to SEQ ID No. 2.

6. A polypeptide according to claim 5 which comprises amino acids 19 to 406 of the amino acid sequence set out in SEQ ID No. 2.

7. A polynucleotide encoding a polypeptide according to any one of the preceding claims.

8. A polynucleotide comprising:

(a) the polynucleotide sequence set out in SEQ ID No. 1, or the complement thereof;

(b) a polynucleotide sequence capable of hybridising to the nucleotide sequence set out in SEQ ID No. 1, or a fragment thereof;

(c) a polynucleotide sequence capable of hybridising to the complement of the polynucleotide sequence set out in SEQ ID No. 1, or a fragment thereof; and/or

(d) a polynucleotide sequence which is degenerate as a result of the genetic code to any of the polynucleotides defined in (a), (b) or (c).

9. A polynucleotide according to claim 8 which:

a. encodes a polypeptide having endo-xylogalacturonase activity, which polynucleotide is:

(1) the coding sequence of SEQ ID No. 1;

(2) a sequence which hybridises selectively to the complement of sequence defined in (1); or

(3) a sequence that is degenerate as a result of the genetic code with respect to a sequence defined in (1) or (2); or

b. is a sequence complementary to a polynucleotide defined in (a).

10. An isolated polynucleotide according to claim 7, 8 or 9 obtainable from a fungus.

11 A polynucleotide according to claim 10 wherein the fungus is of the genus *Aspergillus*.

12. A polynucleotide probe which comprises a fragment of at least 15 nucleotides of a polynucleotide as defined in any of claims 7 to 11.

13. A vector comprising a polynucleotide as defined in any one of claims 7 to 12.

14. An expression vector comprising a polynucleotide as defined in any one of claims 7 to 11 operably linked to one or more regulatory sequences capable of directing expression of the polynucleotide in a host cell.

15. A host cell transformed or transfected with, comprising or incorporating a vector according to any one of claims 13 to 14.

16. A host cell comprising or harbouring a polynucleotide according to any one of claims 7 to 11 wherein the polynucleotide is heterologous to the genome of the host cell.

17. A host cell according to claim 15 or claim 16 which is a yeast cell.

18. A method for producing a polypeptide according to any one of claims 1 to 6 which comprises incubating or culturing a host cell according to any one of claims 15 to 17 under conditions which allow the expression of the polypeptide, and optionally purifying the polypeptide.

19. A host cell comprising or expressing a polypeptide according to any one of claims 1 to 6 wherein the polypeptide is heterologous to the host cell.

20. A composition comprising a polypeptide according to any one of claims 1 to 6

21. A composition according to claim 20 which further comprises a polypeptide having endo-arabinanase, rhamnogalacturonase or polygalacturonase activity.

22. A method of treating a plant material, the method comprising contacting the plant material with a polypeptide according to any one of claims 1 to 6 or a composition according to claim 20 or claim 21.

23. A method according to claim 22 wherein the treatment comprises degrading or modifying pectin in the plant material.

24 A method according to claim 22 for degrading or modifying plant cell walls.

25. A method according to claim 22 or 23 wherein the treatment comprises endo-type cleaving of xylogalacturonan subunits of a pectin component of the material.

26. A method according to any of claims 22 to 24 wherein the material comprises a plant, plant pulp, plant extract or an edible foodstuff or ingredient therefor.

27. A method according to claim 26 wherein the material is fruit or vegetable pulp, juice or extract.

28. A processed plant material obtainable by contacting a plant material with a polypeptide according to any one of claims 1 to 6 or a composition according to claims 20 or claim 21, or which results from a method according to any of claim 22 to 26.

29. A processed plant material according to claim 27 which is a fruit or vegetable juice.

30. A method for reducing the viscosity of a plant material, the method comprising contacting the plant material with a polypeptide according to any one of claims 1 to 6 or a composition according to claim 20 or claim 21 in an amount and under conditions effective to degrade pectin contained in the material.

31. Use of a polypeptide according to any one of claims 1 to 6 or a composition according to claim 20 or claim 21 in a method of treating plant material.

32. Use according to claim 31 wherein the treatment comprises endo-type cleaving xylogalacturonan substituents of pectin in the plant material.

33. Use of a polypeptide according to any one of claims 1 to 6 or a composition according to claim 20 or claim 21 in a method of processing plant pulp, juice or extract which method comprises incubating the pulp, juice or extract with the polypeptide or composition to at least partially degrade pectin.

34. An (animal) feed or foodstuff comprising a polypeptide according to any one of the claims 1 to 6.

35. A composition comprising (optionally saponified) gum tragacanth (sGT) treated with a strong acid.

36 An assay for identifying or detecting a polypeptide having pectin degrading activity, the assay comprising:

a providing, as a substrate for a candidate compound, (optionally saponified) gum tragacanth treated with a strong acid (sGT/TFA); and

b contacting the sGT/TFA with the candidate compound and detecting whether any reducing carbohydrates are produced.

37. An assay according to claim 35 wherein the amount of reducing carbohydrates is measured and optionally compared to the amount of the carbohydrates produced in a control with the absence of the candidate compound.

38. An assay according to claim 35 or 36 which comprises measuring the amount of Cu(II) reduced to Cu(I) by the carbohydrates, optionally by contact with bicinchoninic acid (BCA) and determining the amount of BCA-Cu(I) complex formed.

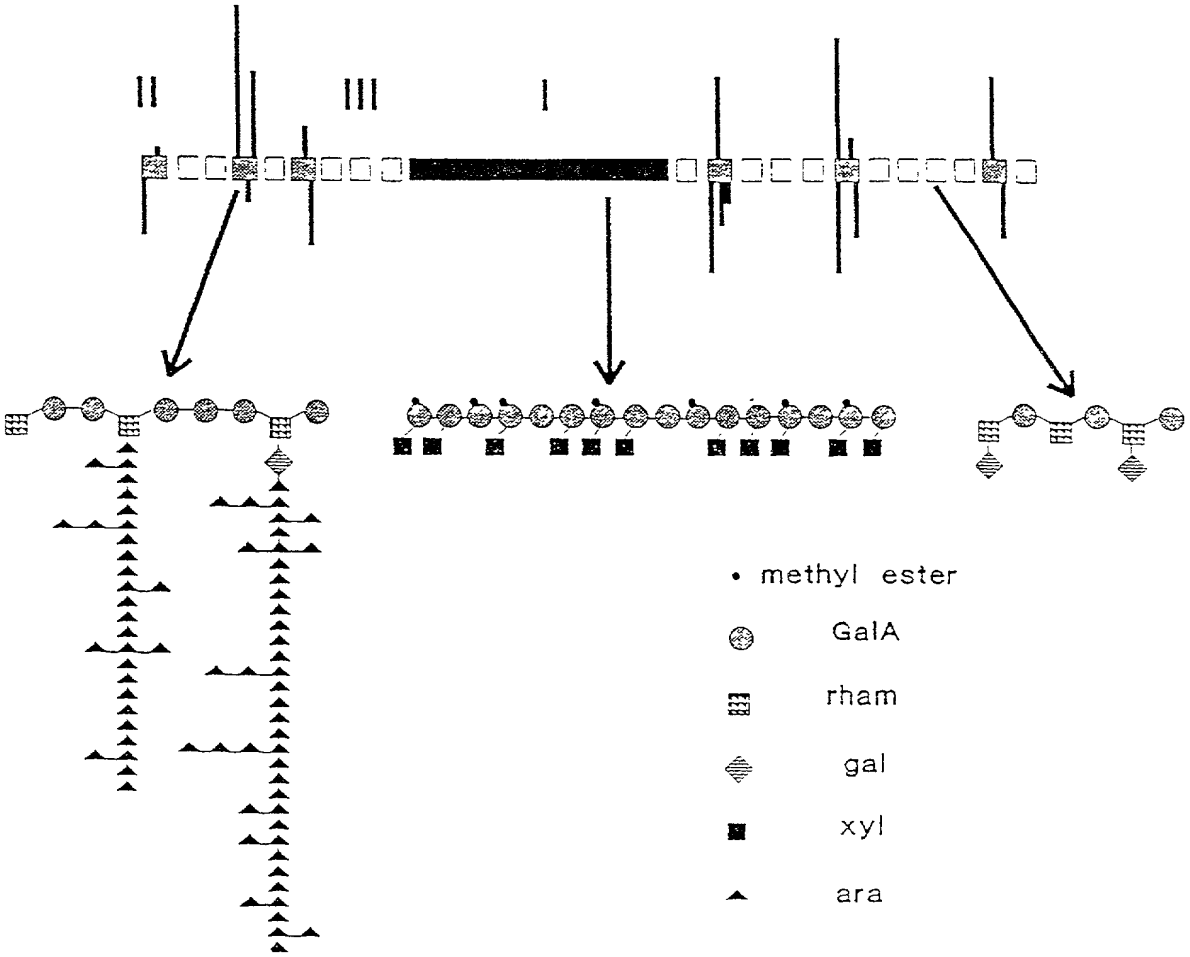


FIGURE 1

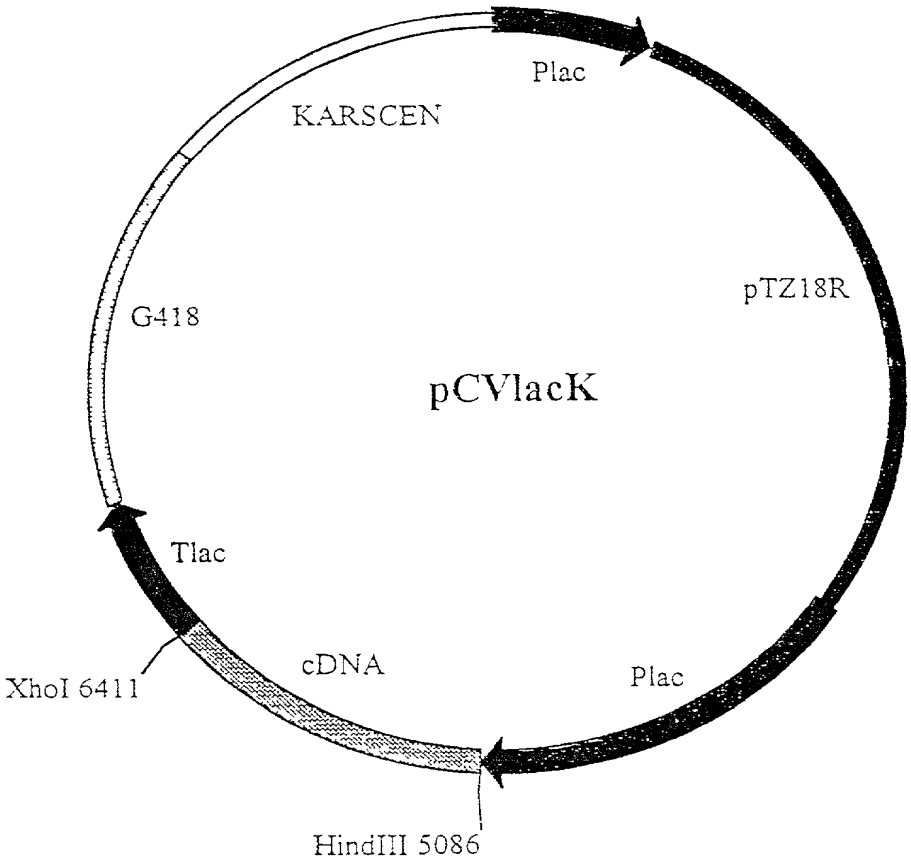


FIGURE 2

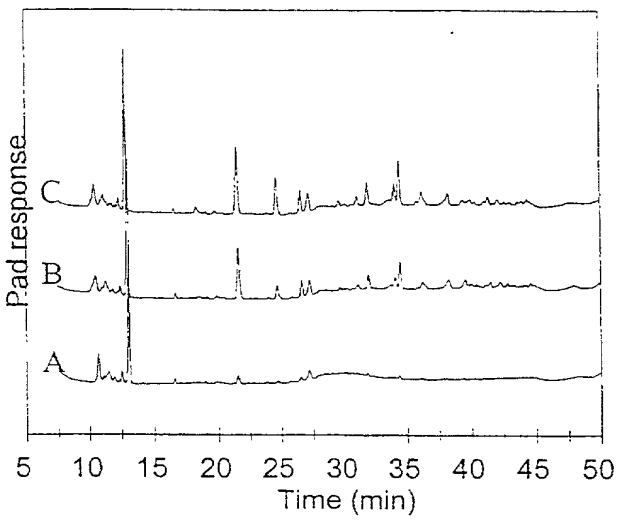


FIGURE 3

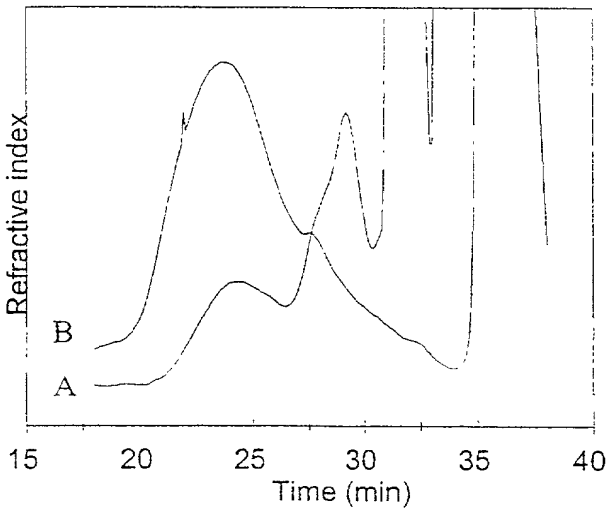


FIGURE 4

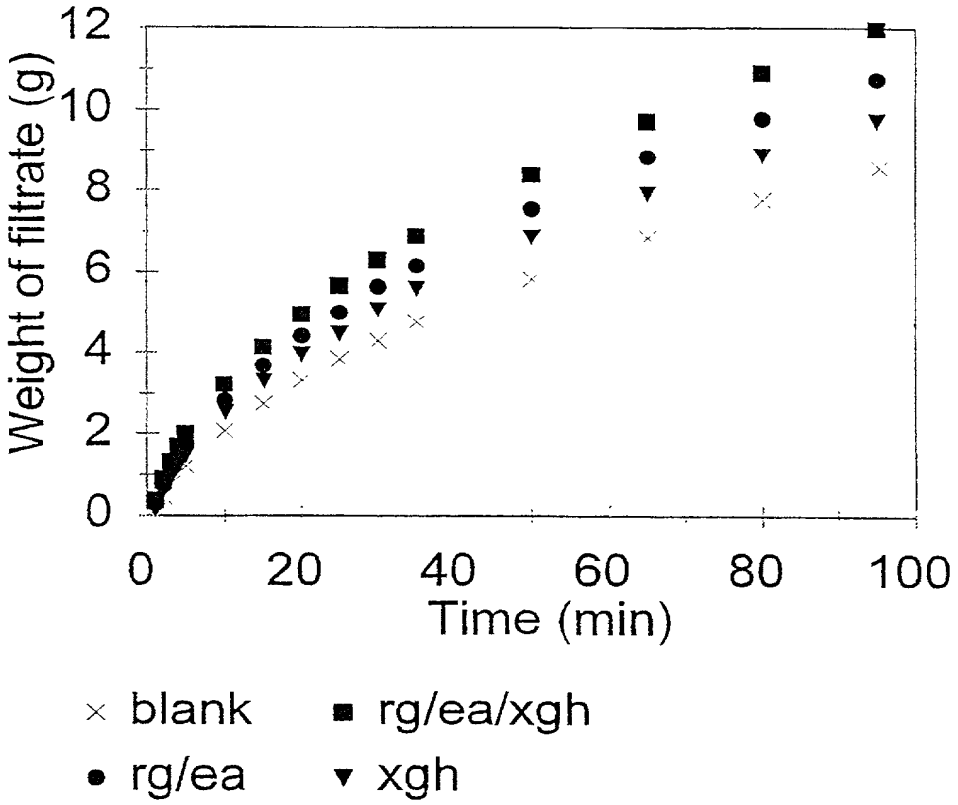


FIGURE 5

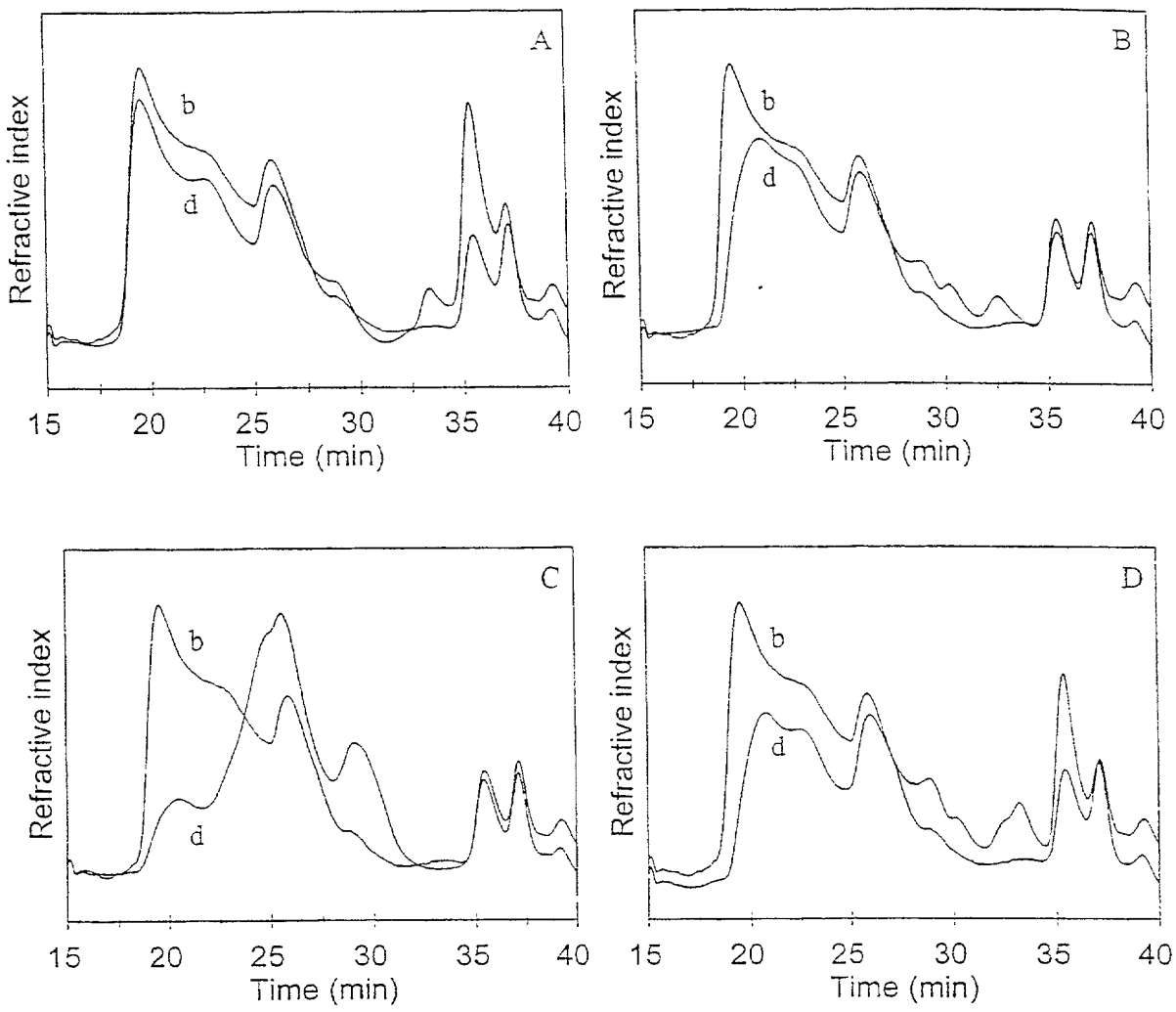


FIGURE 6A to 6D

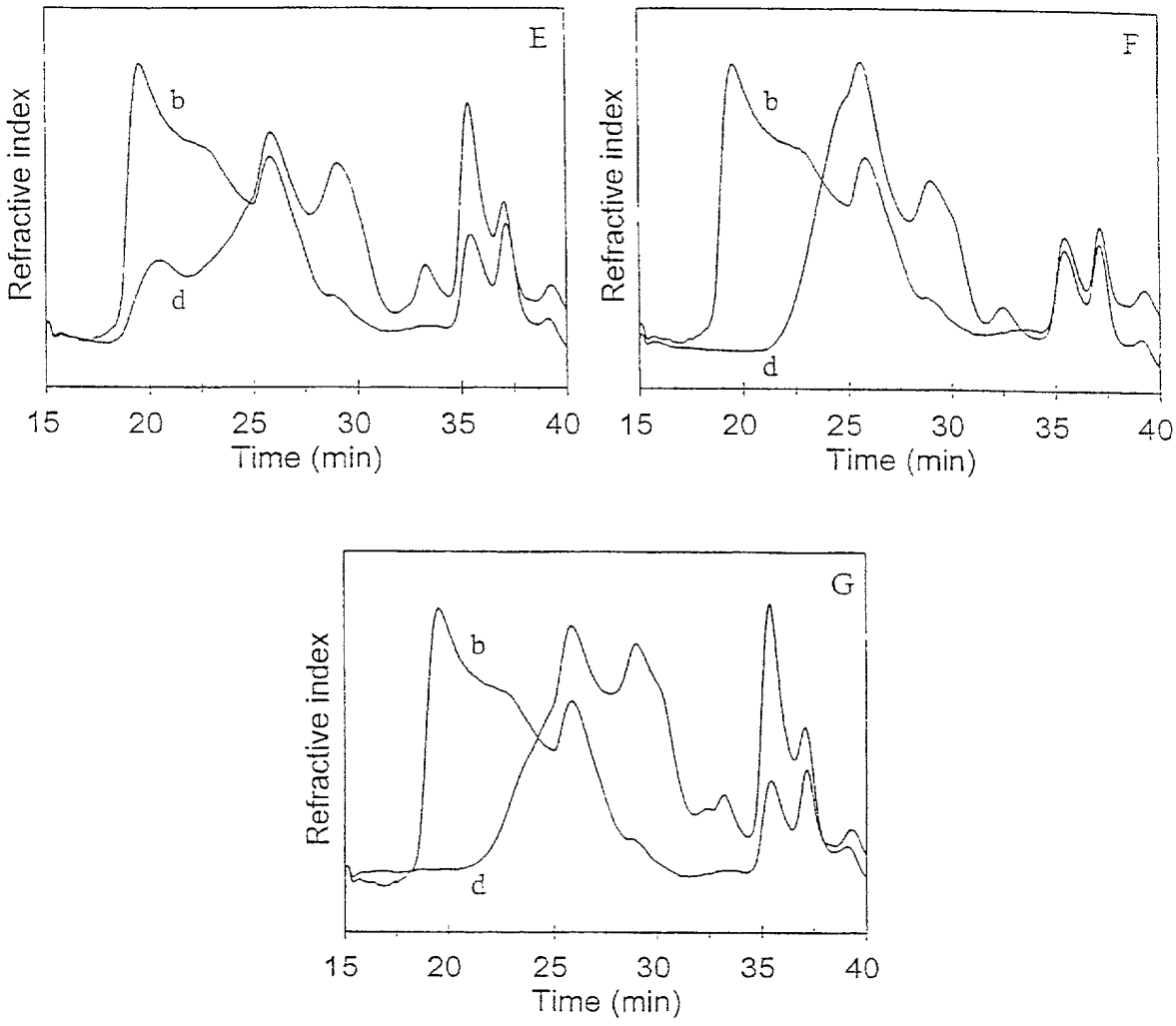


FIGURE 6E to 6G

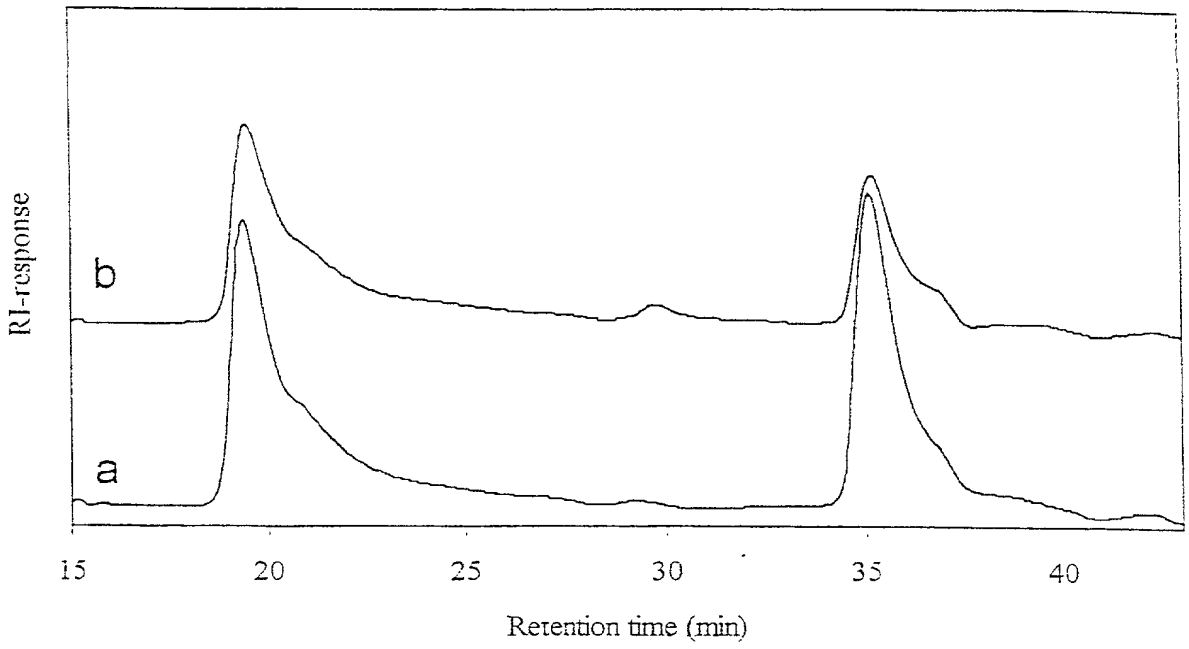


FIGURE 7

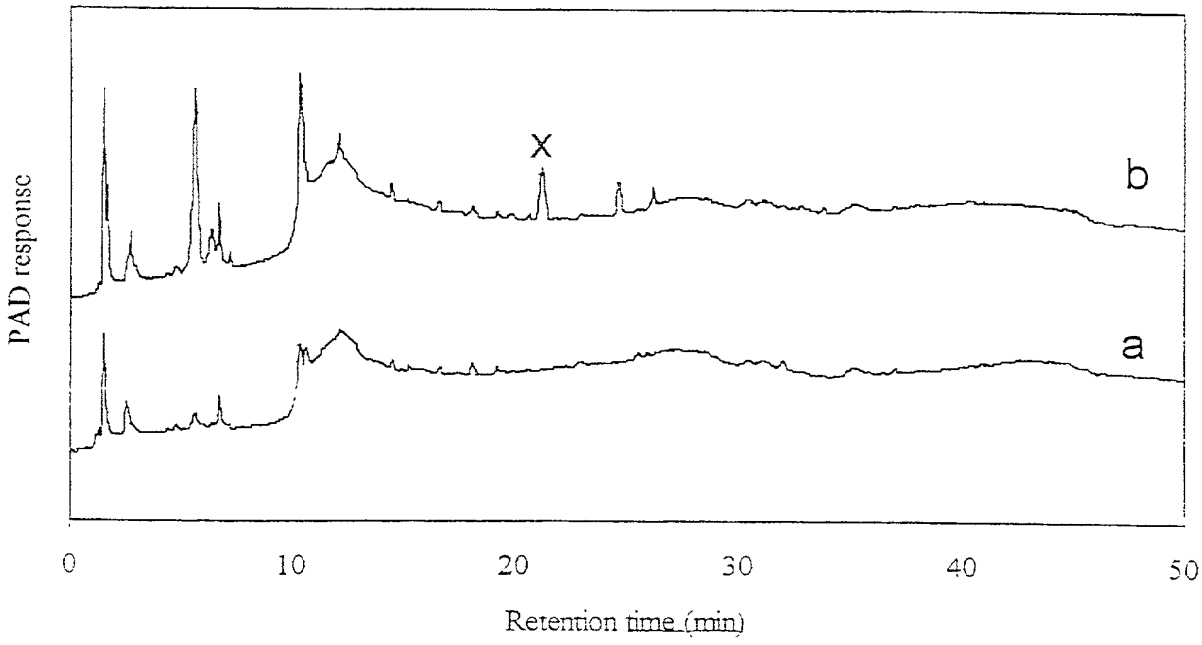


FIGURE 8

```
                          I                                         II                                      III
Atub-PgaII   173  TLGGH...A. .V.N.VG.N. IKPW.H.Q.. .L.IN-.GEN IWFTSGT.I. G..L.I...
Anig-PgaII   173  TQGGH...A. .V.N.VG.N. IKPW.H.Q.. .L.VN-.GEN IWFTGGT.I. G..L.I...
Anig-PgaI    179  ..GGH..... ..SE..G.Y. SGA..K.Q.. .I.IN-.GES ISFTGGT.S. G..L.I..
Anig-PgaC    193  .DLAA..... .......I.. .GAEIY.Q.. ...IN-.GEN IYFSASV.S. G..L.I..
Atub-PgaX    216  ..EAK....W .TYR.NNIV. QNSVIN.G.. ..S...NSTN IL.QNLH.N. ........

Atub-XghA    200  DNPPKNTDGF DIGESTYVTI TEVTVVNDDD CVAFKPSSNY VTVDTISCTG SHGISVGS

 Aac-RhgA    188  GGNEGGL..I .VWG.N-IWV HD.E.T.K.E ..TV.SPA.N IL.ES.Y.NW .G.CAM..
Anig-RhgA    189  GGNSGGL..I .VWG.N-IWV HD.E.T.K.E ..TV.GPA.N IL.ES.Y.NW .G.CAM..
Anig-RhgB    192  GGNHGGL..I .VWSNN-IWV HD.E.T.K.E .VTVKSP.KN ILIES.Y.NW .G.CGM..

                                                 IV
Atub-PgaII   232  GR.N--NV.. .VTIEHS.VS ..EN.VR... VSGATGSVSE ITYSNIVMSG IS.YGVVIQ
Anig-PgaII   232  DR.N--NV.. .VTIEHS.VS ..EN.VR... ISGATGSVSE ITYSNIVMSG IS.YGVVIQ
Anig-PgaI    238  GRD.--NT.. .VTISDS.VS .SANGVR... IYKETGDVSE ITYSNIQLSG IT.YGIVIE
Anig-PgaC    252  GRD.--NT.. .VTFYDVNVL K.QQ.IR... IYGDTGSVSE VTYHEIAFSD AT.YGIVIE
Atub-PgaX    276  QYK.EV.I.E .V..YNISMF .ASDM.R..V W.GTPSALSA DLQGGGGSGS VKN--ITYD

Atub-XghA    260  KSSD--DSVK NIYVTGATMI NSTKAAGIKT YPSGGDHGTS TVSNVTFNDF TVDNSDYAF

 Aac-RhgA    247  --A.--TD.T D.VYRNVYTW S.NQMYM..S -----NG.SG .....LLEN. IGHGNA.SI
Anig-RhgA    248  --A.--TDIT D.LYRNVYTW S.NQMYM..S -----NG.SG ..N.TLLEN. IGRGNR.SI
Anig-RhgB    251  --S.--TN.S D.TYRNIYTW S.NNMML..S -----NG.SG F.E..LLEN. IGHGNA.SI
```

FIGURE 9

US-02905

Case No.: MBHB00-615

DECLARATION AND POWER OF ATTORNEY FOR PATENT APPLICATION

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name.

I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled:

Novel Endo-Xylogalacturonase

the specification of which is attached hereto unless the following space is checked:

☒ was filed on August 9, 2000 as United States Application Serial Number 09/601,852.

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge the duty to disclose information which is material to patentability as defined in 37 CFR § 1.56.

I hereby claim foreign priority benefits under 35 U.S.C. § 119(a)-(d) or § 365(b) of any foreign application(s) for patent or inventor's certificate, or § 365(a) of any PCT international application which designated at least one country other than the United States, listed below and have also identified below, by checking the box, any foreign application for patent or inventor's certificate, or PCT international application having a filing date before that of the application on which priority is claimed.

Prior Foreign Application(s):

	Number	Country	Day/Month/Year Filed
1.	PCT/EP99/00860	PCT	09/02/1999
2.	98300952.3	EP	10/02/1998

I hereby claim the benefit under 35 U.S.C. § 119(e) of any United States provisional application(s) listed below:

	Application Number	Filing Date
1.		
2.		

I hereby claim the benefit under 35 U.S.C. § 120 of any United States application(s), or § 365(c) of any PCT international application designating the United States, listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States or PCT international application in the manner provided by the first paragraph of 35 U.S.C. § 112, I acknowledge the duty to disclose information which is material to patentability as defined in 37 C.F.R. § 1.56 which became available between the filing date of the prior application and the national or PCT international filing date of this application.

	Application Number	Filing Date	Status: patented, pending, abandoned
1.			
2.			

DONNELL BOEHMEN
ILBERT & BERGHOFF
0 SOUTH WACKER DRIVE
ICAGO, ILLINOIS 60606
LEPHONE (312) 913-0001

I hereby appoint the practitioners associated with the Customer Number provided below to prosecute this application and to transact all business in the Patent and Trademark Office connected therewith, and I direct that all correspondence be addressed to that Customer Number.

Customer Number: 020306
Principal attorney or agent: _____
Telephone number: 312-913-0001

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

Full name of first inventor: Petrus J. A. Meeuwsen

Inventor's signature: _____ Date: 30-10-2000
Residence: Arnhem, The Netherlands
Citizenship: Dutch
Post Office Address: Aagje Dekenstraat 60, NL-6836 RN Arnhem, The Netherlands NLX

Full name of second joint inventor: Cecile J. B. Van Der Vlugt-Bergmans

Inventor's signature: C.J.B. v.d. Vlugt-Bergmans Date: 4-11-2000
Residence: Rhenen, The Netherlands
Citizenship: Dutch
Post Office Address: Dahliastraat 13, NL-3911 WB Rhenen, The Netherlands NLX

Full name of third joint inventor: Jean Paul Vincken

Inventor's signature: _____ Date: 11/10/2000
Residence: Renkum, The Netherlands
Citizenship: Dutch
Post Office Address: Willebrordweg 23, NL-6871 ZS Renkum, The Netherlands NLX

Full name of fourth joint inventor: Gerrit Beldman

Inventor's signature: _____ Date: 12/10/2000
Residence: Wageningen, The Netherlands
Citizenship: Dutch
Post Office Address: Ko van Dijkstraat 23, NL-6708 ML Wageningen, The Netherlands NLX

Full name of fifth joint inventor: Alphons G. J. Voragen

Inventor's signature: _____ Date: 12/10/2000
Residence: Wageningen The Netherlands
Citizenship: Dutch
Post Office Address: Sparrenbos 37, NL-6705 BB Wageningen, The Netherlands NLX

McDONNELL BOEHNEN HULBERT & BERGHOFF
300 SOUTH WACKER DRIVE
CHICAGO, ILLINOIS 60606
TELEPHONE (312) 913-0001

Full name of sixth joint inventor: Margareta A. Herweijer

Inventor's signature: _____ Date: 17/11/00
Residence: Den Haag, The Netherlands
Citizenship: Dutch
Post Office Address: Roelofsstraat 43 NL 2596 VK Den Haag, The Netherlands NL

Full name of seventh joint inventor: Albert J.J. Van Ooijen

Inventor's signature: _____ Date: 20/11/00
Residence: Voorburg, The Netherlands
Citizenship: Dutch
Post Office Address: Overburgkade 78, NL-2275, XX Voorburg, The Netherlands NLX

McDONNELL BOEHNEN
HULBERT & BERGHOFF
300 SOUTH WACKER DRIVE
CHICAGO, ILLINOIS 60606
TELEPHONE (312) 913-0001

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:
(A) NAME: Gist-brocades B.V.
(B) STREET: Wateringseweg 1
(C) CITY: Delft
(E) COUNTRY: the Netherlands
(F) POSTAL CODE (ZIP): 2611 XT

(ii) TITLE OF INVENTION: Novel Endo-xylogalacturonase

(iii) NUMBER OF SEQUENCES: 2

(iv) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

(v) CURRENT APPLICATION DATA:
APPLICATION NUMBER: N/A

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1602 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Aspergillus tubigensis

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 98..1318 CTCGAG is XhoI site

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
GCTTGTGTTT CTTAGGAGAA TTATTATTCT TTTGTTATGT TGCGCTTGTA GTTGGAAAAG      60

GTGAAGAGAC AAAGCTTGAA TTCCGAAATC GCTCATC ATG GCG CTA TAT CGT AAC      115
                                         Met Ala Leu Tyr Arg Asn
                                          1               5

CTC TAC CTT CTG GCC AGC CTT GGG CTA AGC AGT GCT GCT CCC TCC AAG      163
Leu Tyr Leu Leu Ala Ser Leu Gly Leu Ser Ser Ala Ala Pro Ser Lys
             10                  15                  20
```

```
GTC CAG CGA GCC CCG GAT TCT TCC ATT CAT GCT CGC GCT GTC TGT ACC      211
Val Gln Arg Ala Pro Asp Ser Ser Ile His Ala Arg Ala Val Cys Thr
         25                  30                  35

CCG ACC GCA GGA GGC GAT TCG TCC ACC GAC GAT GTC CCC GCC ATC ACC      259
Pro Thr Ala Gly Gly Asp Ser Ser Thr Asp Asp Val Pro Ala Ile Thr
     40                  45                  50

GAG GCC CTC AGC TCG TGC GGA AAT GGT GGC ACC ATC GTC TTC CCC GAG      307
Glu Ala Leu Ser Ser Cys Gly Asn Gly Gly Thr Ile Val Phe Pro Glu
 55                  60                  65                  70

GGC AGC ACC TAC TAC CTC AAC AGT GTG CTG GAC TTG GGC AGC TGC AGT      355
Gly Ser Thr Tyr Tyr Leu Asn Ser Val Leu Asp Leu Gly Ser Cys Ser
                 75                  80                  85

GAT TGC GAC ATC CAG GTG GAA GGT CTT CTG AAG TTC GCC AGC GAT ACC      403
Asp Cys Asp Ile Gln Val Glu Gly Leu Leu Lys Phe Ala Ser Asp Thr
             90                  95                 100

GAT TAC TGG AGC GGT CGC ACT GCC ATG ATC AGT GTT TCC AAT GTA GAT      451
Asp Tyr Trp Ser Gly Arg Thr Ala Met Ile Ser Val Ser Asn Val Asp
         105                 110                 115

GGT TTG AAG CTG CGC TCA TTG ACT GGA TCT GGT GTC ATT GAT GGC AAT      499
Gly Leu Lys Leu Arg Ser Leu Thr Gly Ser Gly Val Ile Asp Gly Asn
     120                 125                 130

GGC CAG GAT GCG TGG GAT CTC TTT GCT TCG GAC AGT AGT TAC TCA CGC      547
Gly Gln Asp Ala Trp Asp Leu Phe Ala Ser Asp Ser Ser Tyr Ser Arg
135                 140                 145                 150

CCG ACG CTC TTG TAC ATC ACT GGC GGC AGC AAC CTA GAA ATC TCC GGG      595
Pro Thr Leu Leu Tyr Ile Thr Gly Gly Ser Asn Leu Glu Ile Ser Gly
                155                 160                 165

CTG CGT CAA AAG AAT CCA CCT AAC GTG TTC AAC TCG GTC AAG GGT GGC      643
Leu Arg Gln Lys Asn Pro Pro Asn Val Phe Asn Ser Val Lys Gly Gly
            170                 175                 180

GCC ACT AAT GTC GTC TTC TCC AAC CTG AAG ATG GAT GCC AAC TCC AAG      691
Ala Thr Asn Val Val Phe Ser Asn Leu Lys Met Asp Ala Asn Ser Lys
        185                 190                 195

TCG GAC AAT CCG CCC AAG AAC ACT GAT GGG TTC GAC ATT GGC GAG AGT      739
Ser Asp Asn Pro Pro Lys Asn Thr Asp Gly Phe Asp Ile Gly Glu Ser
    200                 205                 210

ACC TAT GTG ACC ATC ACC GAG GTC ACC GTA GTC AAC GAT GAC GAC TGT      787
Thr Tyr Val Thr Ile Thr Glu Val Thr Val Val Asn Asp Asp Asp Cys
215                 220                 225                 230

GTC GCC TTC AAG CCC AGT TCC AAC TAC GTG ACA GTG GAC ACG ATC AGC      835
Val Ala Phe Lys Pro Ser Ser Asn Tyr Val Thr Val Asp Thr Ile Ser
                235                 240                 245

TGC ACC GGC TCC CAT GGA ATT TCC GTG GGA TCA TTA GGA AAG TCG AGC      883
Cys Thr Gly Ser His Gly Ile Ser Val Gly Ser Leu Gly Lys Ser Ser
```

```
                250                 255                 260

GAC GAC TCG GTC AAG AAC ATT TAT GTC ACG GGC GCA ACT ATG ATC AAC      931
Asp Asp Ser Val Lys Asn Ile Tyr Val Thr Gly Ala Thr Met Ile Asn
        265                 270                 275

TCC ACC AAA GCC GCC GGG ATC AAG ACT TAT CCG AGT GGA GGC GAC CAC      979
Ser Thr Lys Ala Ala Gly Ile Lys Thr Tyr Pro Ser Gly Gly Asp His
    280                 285                 290

GGT ACC TCC ACG GTC AGC AAT GTG ACC TTC AAC GAT TTC ACT GTG GAC     1027
Gly Thr Ser Thr Val Ser Asn Val Thr Phe Asn Asp Phe Thr Val Asp
295                 300                 305                 310

AAC TCC GAC TAT GCC TTC CAG ATC CAG AGC TGC TAT GGC GAG GAC GAT     1075
Asn Ser Asp Tyr Ala Phe Gln Ile Gln Ser Cys Tyr Gly Glu Asp Asp
                315                 320                 325

GAC TAT TGC GAG GAA AAC CCG GGC AAC GCC AAA CTG ACT GAT ATA GTC     1123
Asp Tyr Cys Glu Glu Asn Pro Gly Asn Ala Lys Leu Thr Asp Ile Val
            330                 335                 340

GTG TCA AGC TTC AGT GGG ACA ACC AGT GAC AAG TAC GAT CCG GTC GTG     1171
Val Ser Ser Phe Ser Gly Thr Thr Ser Asp Lys Tyr Asp Pro Val Val
        345                 350                 355

GCC AAC CTC GAC TGC GGT GCG GAT GGA ACT TGT GGC ATC TCC ATC AGT     1219
Ala Asn Leu Asp Cys Gly Ala Asp Gly Thr Cys Gly Ile Ser Ile Ser
    360                 365                 370

GGG TTC GAT GTC AAG GCG CCA TCG GGC AAG TCT GAA GTG TTG TGC GCC     1267
Gly Phe Asp Val Lys Ala Pro Ser Gly Lys Ser Glu Val Leu Cys Ala
375                 380                 385                 390

AAC ACC CCG TCT GAT TTG GGC GTC ACT TGC ACT TCG GGG GCT TCG GGC     1315
Asn Thr Pro Ser Asp Leu Gly Val Thr Cys Thr Ser Gly Ala Ser Gly
                395                 400                 405

TAAATAGCTT TGGCCGGGTT GCTTTCTGAA TCCACTGAGT GGAGGTCTTC TTCGGGTTTG    1375

ATATTTTGTA TGGTCGTGTG TATAGCAGAA TGTGACAATA GAATTAGTGA AATTGCCATT    1435

GTTTTCGAAA GACAAAAAAA AAAAAAAAAA AAAAAAAAAA ACTCGAGAAT TTATACTTAG    1495

ATAAGTATGT ACTTACAGGT ATATTTCTAT GAGATACTGA TGTATACATG CATGATAATA    1555

TTTAAACGGT TATTAGTGCC GATTGTCTTG TGCGATAATG ACGTTCC                  1602
```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 406 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Ala Leu Tyr Arg Asn Leu Tyr Leu Leu Ala Ser Leu Gly Leu Ser
1 5 10 15

Ser Ala Ala Pro Ser Lys Val Gln Arg Ala Pro Asp Ser Ser Ile His
20 25 30

Ala Arg Ala Val Cys Thr Pro Thr Ala Gly Gly Asp Ser Ser Thr Asp
35 40 45

Asp Val Pro Ala Ile Thr Glu Ala Leu Ser Ser Cys Gly Asn Gly Gly
50 55 60

Thr Ile Val Phe Pro Glu Gly Ser Thr Tyr Tyr Leu Asn Ser Val Leu
65 70 75 80

Asp Leu Gly Ser Cys Ser Asp Cys Asp Ile Gln Val Glu Gly Leu Leu
85 90 95

Lys Phe Ala Ser Asp Thr Asp Tyr Trp Ser Gly Arg Thr Ala Met Ile
100 105 110

Ser Val Ser Asn Val Asp Gly Leu Lys Leu Arg Ser Leu Thr Gly Ser
115 120 125

Gly Val Ile Asp Gly Asn Gly Gln Asp Ala Trp Asp Leu Phe Ala Ser
130 135 140

Asp Ser Ser Tyr Ser Arg Pro Thr Leu Leu Tyr Ile Thr Gly Gly Ser
145 150 155 160

Asn Leu Glu Ile Ser Gly Leu Arg Gln Lys Asn Pro Pro Asn Val Phe
165 170 175

Asn Ser Val Lys Gly Gly Ala Thr Asn Val Val Phe Ser Asn Leu Lys
180 185 190

Met Asp Ala Asn Ser Lys Ser Asp Asn Pro Pro Lys Asn Thr Asp Gly
195 200 205

Phe Asp Ile Gly Glu Ser Thr Tyr Val Thr Ile Thr Glu Val Thr Val
210 215 220

Val Asn Asp Asp Asp Cys Val Ala Phe Lys Pro Ser Ser Asn Tyr Val
225 230 235 240

Thr Val Asp Thr Ile Ser Cys Thr Gly Ser His Gly Ile Ser Val Gly
245 250 255

Ser Leu Gly Lys Ser Ser Asp Asp Ser Val Lys Asn Ile Tyr Val Thr
260 265 270

Gly Ala Thr Met Ile Asn Ser Thr Lys Ala Ala Gly Ile Lys Thr Tyr
275 280 285

Pro Ser Gly Gly Asp His Gly Thr Ser Thr Val Ser Asn Val Thr Phe
290 295 300

```
Asn Asp Phe Thr Val Asp Asn Ser Asp Tyr Ala Phe Gln Ile Gln Ser
305                 310                 315                 320

Cys Tyr Gly Glu Asp Asp Asp Tyr Cys Glu Glu Asn Pro Gly Asn Ala
                325                 330                 335

Lys Leu Thr Asp Ile Val Val Ser Ser Phe Ser Gly Thr Thr Ser Asp
            340                 345                 350

Lys Tyr Asp Pro Val Val Ala Asn Leu Asp Cys Gly Ala Asp Gly Thr
        355                 360                 365

Cys Gly Ile Ser Ile Ser Gly Phe Asp Val Lys Ala Pro Ser Gly Lys
    370                 375                 380

Ser Glu Val Leu Cys Ala Asn Thr Pro Ser Asp Leu Gly Val Thr Cys
385                 390                 395                 400

Thr Ser Gly Ala Ser Gly
                405
```